

Le Web 2.0 pour la veille et la recherche d'information

Exploitez toutes les ressources du Web Social

Juin 2007

White Paper

Avertissement

Ce document a été réalisé par la société Digimind.

Le contenu de ce document est protégé par le droit d'auteur. Il peut être reproduit en partie à la condition de citer Digimind comme auteur et d'indiquer l'adresse <http://www.digimind.com>. Pour toute information complémentaire, vous pouvez contacter Digimind par mail à l'adresse contact@digimind.com ou par téléphone au 01 53 34 08 08.

Sommaire

AVERTISSEMENT	1
AVANT PROPOS	4
Web 2.0 ?	4
QU'EST-CE QUE LE WEB 2.0 ?	5
Préambule	5
Définition du web 2.0	6
Les 7 principes du Web 2.0	8
Les évolutions par rapport au Web 1.0 : le rôle de l'internaute	9
GLOSSAIRE DU WEB 2.0	12
LES TECHNOLOGIES DU WEB 2.0	15
LES FAMILLES D'APPLICATIONS DU WEB 2.0	18
Typologie	18
Comment identifier un site Web 2.0 ?	23
LES APPLICATIONS WEB 2.0 POUR LA VEILLE ET LA RECHERCHE D'INFORMATIONS	24
Les types d'applications	24
I. Partager et collaborer	24
a. Les Bookmarks sociaux	24
b. Le Multimédia : photos et vidéos	30
c. Partage de cv et réseaux sociaux.....	35
d. Les sites d'actualités collaboratives : recommandation et notation.....	38
II. Rechercher, Collecter	42
a. La recherche de blogs et de flux RSS.....	42
b. La recherche par tags.....	48
c. La recherche de personnes et d'informations entreprises	54
d. Les moteurs de recherche "humains"	56
e. les moteurs de recherche "collaboratifs"	59
f. La verticalisation par la personnalisation du périmètre de recherche.....	62
g. L'enrichissement de bases de données façon Web2.0	65
h. Conserver et exploiter l'historique de vos recherches et navigations.....	67
i. Les "mash-up"	69
III. Diffuser.....	73

a. Les Blogs et le flux RSS	73
b. Les Wikis	74
IV. Surveiller	78
a. La détection du flux RSS	78
b. La mise en surveillance.....	79
c. Pourquoi une veille via le RSS ?.....	81
V. Carte des outils et ressources Web 2.0.....	82
Exemple d'une recherche et veille via le Web 2.0	84
LES LIMITES DU WEB 2.0	86
La participation.....	86
I. Une participation hétérogène	86
Une minorité participative... ..	86
Wikipedia, un faux modèle collaboratif ?.....	87
II. Une participation d'"amateurs" ?.....	88
Le web 1.0 reste nécessaire	89
Le mail résiste	89
Une population souvent isolée et non représentative	90
Des domaines mal couverts	91
CONCLUSION	92
A PROPOS DES AUTEURS	96
WEBOGRAPHIE BIBLIOGRAPHIE	98
Sites web, wikis et blogs.....	98
Articles	101
Ouvrages	102
DANS LA MEME COLLECTION.....	104
NOTES.....	105

Avant propos

Web 2.0 ?

La notion de Web 2.0 est très souvent employée mais, finalement, peu de personnes sont capables de l'expliquer précisément.

L'objectif de ce document est de "démocratiser" le concept de Web 2.0 auprès des "Knowledge Workers". Il s'adresse aux différents collaborateurs de l'entreprise qui n'utilisent pas encore ou peu les différentes applications regroupées sous ce concept. Les sensibiliser à ces outils pourra améliorer la pertinence de leur recherche et de leur veille informationnelle.

Qu'est-ce que le Web 2.0 ?

PREAMBULE

Pour certains, le Web 2.0 n'est qu'un gadget, un pur phénomène de mode qui ne mérite pas que l'on s'y attarde. Pourtant, une évolution importante a bien eu lieu sur le web, quelque soit le nom que l'on lui donne.

Ainsi, en 2006, un metteur en scène a choisi de diffuser son long métrage, *Autumn*¹, gratuitement, et ce uniquement sur internet, via le service *Google Video*². Il n'avait pas trouvé de distributeur capable d'exploiter son film dans plus de 30 salles. En mai 2007, plus de 920 000 internautes avaient déjà vu *Autumn* sur le service de partage de vidéos de Google.

Fin 2005, pour le compte de la prestigieuse revue scientifique *Nature*, des experts ont comparé des documents publiés sur Wikipedia, l'encyclopédie collaborative en ligne, réalisée par des internautes volontaires et bénévoles, avec ceux de la célèbre *Encyclopaedia Britannica*. Résultat : la qualité des deux sources est assez proche. Cette étude a abouti à une formidable controverse : comment une encyclopédie réalisée par des "non-spécialistes" peut-elle remettre en cause les ouvrages classiques ?

Il se passe donc bien quelque chose de nouveau via ce web nouvelle génération, où l'internaute devient actif et collaboratif. Mais tellement d'acteurs et d'applications se réclament de ce web 2 que l'on a vite l'impression de naviguer en *Terra Incognita*. Certains n'hésitent d'ailleurs pas à représenter ce nouveau web via une carte d'explorateur.³

Aussi, à travers ce document, plongeons dans le Web 2.0 pour mieux l'exploiter lors de recherche d'information et de veilles.

Plusieurs auteurs s'accordent sur l'idée selon laquelle le web 2.0 ne constitue pas une rupture technologique ni ... sociale. Il s'agirait plutôt d'un retour aux sources, d'une renaissance du web tel qu'il a été conçu à l'origine, les utilisateurs s'étant entre temps appropriés les modes de création et de diffusion. Pour Paul Graham⁵, pionnier de l'Internet et fondateur de la société Viaweb en 1995, *"le web 2.0 est le web comme il a été conçu pour être utilisé. Les tendances que nous distinguons sont simplement la nature inhérente du web qui émerge des mauvaises pratiques qui lui ont été imposées pendant la bulle Internet"*.

Le web créé au début des années 90 par Tim Berners Lee comportait déjà en germe bien des éléments du web 2.0 : mutualisation des connaissances, échange, travail collaboratif, interactivité, "intelligence collective"... Ce n'est sans doute pas un hasard si Dale Dougherty, l'un des pères fondateurs de la formule "2.0", fut lui aussi l'un des pionniers du web, et notamment le fondateur de GNN, le premier portail financé par la publicité.

Naissance du concept

C'est au cours de l'été 2004 que le terme de "web 2.0" apparaît pour la première fois : lors d'une séance de brainstorming, Dale Dougherty, co-fondateur de la société d'édition O'Reilly, et Craig Cline de MediaLive, réfléchissent à des thèmes émergents en vue d'une conférence à venir. Le journaliste John Battelle, co-fondateur du magazine Wired, se joint un peu plus tard au duo et la première conférence Web 2.0 a lieu en octobre 2004. L'expression 2.0 est utilisée à ce propos pour marquer l'émergence d'une nouvelle étape dans la courte histoire du web, étape caractérisée par un changement des règles et une modification des modèles de revenus. Le volet social sera développé plus tardivement par Tim O'Reilly lors de la deuxième conférence Web 2.0 en octobre 2005.

Les nouveaux produits et services associés au web 2.0 vont ainsi repositionner l'internaute au cœur du système. Des techniques de publications simples et souples comme les blogs, les plates-formes collaboratives de type Wiki offrent des possibilités accrues de production, diffusion et consommation des contenus. L'internaute devient ainsi co-créateur de nouvelles applications en ligne, initiant souvent de façon collective ou communautaire de nouvelles formes de relations numériques. Selon l'expression d'Hubert Guillaud⁶, le web devient non plus une aventure technologique, mais bien une aventure humaine.

LES 7 PRINCIPES DU WEB 2.0

Le texte de Tim O'Reilly, paru en 2005 ⁷, sous le titre "Qu'est ce que le web 2.0 ?" dégage sept principes clés du Web 2.0, en voici un petit résumé :

1. Le web vu comme une plate-forme de services. On passe d'une collection de sites web à une plateforme informatique à part entière, fournissant des applications web aux utilisateurs.
2. Considérer les internautes comme co-développeurs des applications. On passe ainsi de la notion de "logiciel produit" à celle de "logiciel service".
3. Le service s'améliore quand le nombre d'utilisateurs augmente. Le Web 2.0 met à profit l'effet de la "longue traîne" (ou *long tail* en anglais), popularisé par le magazine *Wired* : les produits qui sont l'objet d'une faible demande, ou qui n'ont qu'un faible volume de vente, peuvent collectivement représenter une part de marché égale ou supérieure à celle des best-sellers, si les canaux de distribution peuvent proposer assez de choix. Les utilisateurs du Web 2.0 disposent de données uniques, difficiles à recréer, et dont la richesse s'accroît avec l'augmentation du nombre des utilisateurs.
4. La richesse est dans les données : Toutes les applications web d'importance sont liées à une base de données spécialisée. O'Reilly envisage un mouvement "des données libres" s'opposant peu à peu à l'univers des données propriétaires.
5. Tirer parti de l'intelligence collective : c'est le principe même adopté par l'encyclopédie en ligne Wikipedia et d'autres sites similaires. Pour Tim O'Reilly, "l'implication des utilisateurs dans le réseau est le facteur-clé pour la suprématie sur le marché".
6. Mettre en place des interfaces souples et légères fondées sur les nouveaux standards et protocoles du Web. Le logiciel se libère du PC. Une autre des caractéristiques du web 2.0 est le fait qu'il n'est plus limité à la plate-forme PC, mais vise notamment les "objets nomades", téléphone portable, PDA ou lecteur portatif audio.



LES EVOLUTIONS PAR RAPPORT AU WEB 1.0 : LE ROLE DE L'INTERNAUTE

En quoi le Web 2.0 se distingue-t-il de la vague précédente ? On a vu que l'une des caractéristiques majeures du phénomène est l'appropriation des contenus et des modes de diffusion par les internautes eux-mêmes, via des applications souples et légères.

Paul Graham évoque le principe de "démocratie" du Web 2.0, c'est-à-dire la superposition au réseau technique d'un réseau humain qui participe à la création et à l'enrichissement de contenus et services. L'internaute deviendrait ainsi "consomm'acteur" : à lui de produire, publier, diffuser, indexer, partager des contenus de tous types.

	Web 1.0	Web 2.0
Leaders du web	Entreprises, marchands	Internautes
Profil de l'internaute	Passif	Actif
Interactivité perçue	Sélection et lecture d'information	Sélection, lecture et publication de données
Unité de recherche	Mot-clé	Tag

▲ Du Web 1 vers le Web 2. Source ZDNet ⁸

Pour autant, les internautes vont-ils devenir tous acteurs de ce processus ? Selon un article de Charles Arthur dans *The Guardian* ⁹, pour un groupe de 100 personnes, si l'une crée du contenu, 10 seulement vont interagir en apportant commentaires et améliorations, les autres se contentant d'une consommation passive. Ce constat, obtenu via des retours d'expérience sur des sites collaboratifs, illustre la "règle des 1/10/89 %", selon laquelle plus de 80% des contenus produits proviennent d'un nombre très restreint d'utilisateurs.

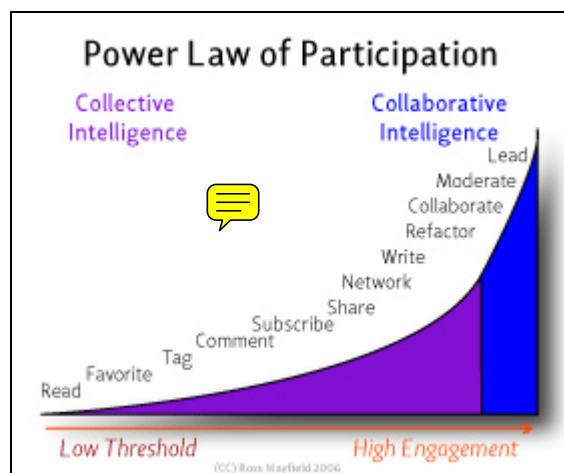
Tendance confirmée par les chiffres présentés lors de la *Web 2.0 Expo* en avril 2007 et repris dans *Internetactu* ¹⁰ : selon la société de mesure d'audience *Hitwise* ¹¹, "0,16 % des internautes qui passent sur *YouTube* y placent des films - alors qu'ils étaient encore 0,5 % à le faire en juillet 2006 ; seulement 0,2% des visiteurs de *Flickr* y publient des photos. La seule exception à cette faible participation provient

de *Wikipedia*, où 4,6 % des visiteurs contribuent à l'enrichissement de l'encyclopédie".

Jimmy Wales, fondateur de *Wikipedia*, estime qu'un nombre restreint de contributeurs et éditeurs font l'essentiel du travail, alors qu'un nombre considérable de lecteurs apportent peu de modifications. Le "wikipédien" Aaron Swartz nuance ce point de vue : "un grand nombre de contributeurs occasionnels semblent en effet avoir écrit des textes longs (en un nombre réduit d'interventions), tandis que les "éditeurs" les plus actifs, qui forment le noyau des wikipédiens nettoient, corrigent, mais sans apporter énormément de texte".




Le Web 2.0, et tout particulièrement la blogosphère semble suivre la loi de la participation, édictée par Ross Mayfield, président de la société Social Text, sous le nom de "Power law of participation"¹². Cette "loi", basée sur l'observation des pratiques et des seuils de participation aux dispositifs communautaires, constate que la participation s'amenuise à mesure que les formes de participation se complexifient. Le schéma ci-dessous, qui rappelle celui de la Longue Traîne, distingue l'intelligence collective fondée sur des actions à faible niveau d'exigence (lire, enregistrer un signet, tagger, commenter), de l'intelligence collaborative, caractérisée par des formes de participations plus actives et intenses : écrire, collaborer, modérer...



▲ Power Law of Participation, par Ross Mayfield. CC

Le cabinet *Forrester* détaille encore davantage les niveaux d'implication dans la grille de la participation figurant dans son étude *Social Technographics*¹³ : la grille compte 6 échelons correspondant aux niveaux de participation :

- . les inactifs (52 %)
- . les spectateurs (33 %) : consommateurs de contenus texte ou multimédia (blogs, vidéos mis en ligne par les utilisateurs, podcasts) 
- . les menuisiers (19 %) : utilisateurs de réseaux sociaux
- . les collecteurs (15 %) : utilisent des flux RSS et taggent les pages web
- . les critiques (19 %) : postent des commentaires sur les blogs, des avis et critiques
- . les créateurs (13 %) : publient des contenus en ligne (blogs, vidéos..)

Glossaire du Web 2.0

Pour mieux comprendre les concepts utilisés dans l'univers du Web 2.0, nous vous proposons ce glossaire des termes essentiels :

- **Tags** : mot clef sous forme d' "étiquette" que les utilisateurs peuvent apposer sur un document numérique (texte, photo, vidéo), de façon à en décrire le contenu. Le *tagging* introduit une part de subjectivité par rapport à l'indexation automatique effectuée par les moteurs. Les tags peuvent également être extraits automatiquement d'un texte en fonction de leur fréquence d'occurrence.
- **Nuage de tags** : présente les tags correspondant à un site (ou à plusieurs) en les regroupant sur un même espace et leur attribuant une taille variable suivant la fréquence d'utilisation ou d'occurrence. Cela donne une vision d'ensemble des principales thématiques. Le site *Tagcrowd.com*¹⁴ permet de générer un nuage de tags à partir d'un texte.
- **Folksonomie** : (traduit également parfois par potonomie, peuplonomie...) Ce mot valise, contraction de folk (peuple) et taxonomie, désigne une forme de "classification collaborative décentralisée spontanée" basée non pas sur un vocabulaire contrôlé et standardisé mais sur des termes choisis par les utilisateurs eux-mêmes, intitulés tags. L'objectif étant de faciliter l'indexation des contenus, ainsi que la recherche d'information.
- **Blogs** : journal personnel disponible sur le web. Peut être tenu par un particulier, un chercheur, un journaliste, un salarié d'entreprise ou plus rarement par un groupe de personne (entreprise, collectivité...). Le blog a pour objet de diffuser des billets, généralement liés à l'actualité, et présentés par ordre chronologique, alors qu'un wiki est destiné à la mutualisation des connaissances. Un blog peut présenter des contenus multimédias : photo (photoblog), son ou vidéo (vblog, vlog)...
- **Blogosphère** : la "biosphère des blogs ". Ce mot valise désigne la communauté des blogs, ou encore l'ensemble des auteurs de blogs. La blogosphère est largement interconnectée, les **trackbacks** situés sur la plupart des blogs permettant de lier

deux articles traitant de sujets connexes. L'élite des blogueurs particulièrement influents peut être qualifiée ironiquement de "blogoisie".

- **Blogroll** ou **blogliste** : liste de liens vers d'autres blogs, présentés par l'auteur d'un blog. On peut syndiquer sur une même page des billets venant de ces blogs via les formats RSS.

- **RSS**: l'acronyme signifie selon les versions *Really Simple Syndication* (RSS 2.0), *RDF Site Summary* (RSS 0.9, 1.0 et 1.1) ou encore *Rich Site Summary* (RSS 0.91). Ce format de syndication permet d'extraire automatiquement d'un site web ou d'un blog du contenu régulièrement mis à jour. Un fichier RSS est un simple fichier texte au format XML comportant la description synthétique du contenu. Ce format a été créé en 1999 par Netscape. On peut s'abonner gratuitement à plusieurs flux RSS à partir d'un lecteur on-line ou installé sur le PC.

- **Atom** : Format normalisé de syndication de contenu Web, basé sur le XML, qui permet, la syndication de contenu entre différentes ressources Web. Concurrent de RSS, Le format Atom a été créé dans le but de simplifier les différents formats RSS existants. La différence entre les deux formats réside surtout dans l'emploi de balises XML différentes. Les deux formats sont généralement lus par les agrégateurs.

- **Wiki** : Site web dynamique dont tout visiteur peut modifier les pages à loisir. Le terme vient la langue hawaïenne (wikiwiki = rapide). Un wiki a pour objectif de mutualiser des connaissances et est organisé autour d'un plan de classement. Les contributeurs d'un wiki sont généralement anonymes.

- **Crowdsourcing** : Littéralement "approvisionnement par la foule". Le *crowdsourcing* consiste à utiliser la créativité, l'intelligence et le savoir-faire d'un grand nombre d'internauts pour créer des contenus, participer à l'alimentation d'un site. Cette participation des utilisateurs à la création des contenus représente une forme d'externalisation (outsourcing) et permet de réduire les coûts en se basant sur un système de bénévolat ou de micropaiement.

- **Réseaux sociaux** : de nombreux sites de "social networking" ou réseautage social sont apparus avec la vague du Web 2.0. Ils permettent de mettre en relation des personnes partageant les mêmes centres d'intérêt personnels ou professionnels. Ces réseaux peuvent devenir de véritables communautés en ligne, les plus fameux étant les sites *Friendster* ou *Myspace*, ou bien dans la sphère professionnelle

Linkedin, Viadeo... La plupart des réseaux sont ouverts, certains fonctionnent sur le mode de la cooptation.

- **Podcasts** : issu de la combinaison des termes *iPod* (d'Apple) et *broadcasting*, ce terme désigne le fait de rendre disponible en ligne un fichier audio au format numérique. Ce fichier peut-être téléchargé directement sur un ordinateur ou un périphérique, via un flux de diffusion. Equivalent français : baladodiffusion. L'utilisateur peut ainsi se composer une liste de lecture à la carte. Pour les fichiers vidéo, on emploie le terme de vidéocast.

- **AJAX** (Asynchronous JavaScript And XML) : méthode informatique de développement d'applications Web permettant d'économiser de la bande passante, en ne rechargeant pas une page entière alors que seuls certains éléments ont besoin de l'être, mais en ne rafraîchissant que ces éléments de la page. Cela permet de produire des contenus dynamiques et réactifs en ne nécessitant que peu d'appels au serveur hébergeant la page. AJAX est une combinaison de plusieurs technologies : HTML, CSS, Javascript, XML... De nombreux sites web 2.0 sont développés en AJAX.

- **API** : Interface de programmation permettant à un utilisateur de développer des applications conviviales, De nombreux sites classiques ou "Web 2.0" proposent des API pour intégrer automatiquement leur contenu à d'autres sites.

- **Mashup** ou **Mash up** : application web composite mixant plusieurs sources ou plusieurs contenus pour fournir un nouveau produit ou service. Le terme se traduit en français par remixage, ou mosaïque. Exemple de mashup : la visualisation des résultats des élections présidentielles françaises via Google Earth qui mixe les cartes Google et les chiffres du Ministère de l'Intérieur.

- **Long tail (longue traîne)** : Cette expression est utilisée en statistiques, et peut s'appliquer au marketing. Des produits rares ou qui font l'objet d'une faible demande, peuvent collectivement représenter une part de marché égale ou supérieure à celle des best-sellers, si les canaux de distribution peuvent proposer assez de choix. Le caractère virtuel du web permet de générer une longue série de marchés de niche hétérogènes (exemple : *e-bay*), par opposition aux grands marchés homogènes et grand public du monde réel.

Les technologies du Web 2.0

Mettons d'abord les choses au point : le Web 2.0 n'est pas une (r)évolution ou un concept de dimension technologique.

Ses applications s'appuient sur des technologies déjà existantes répondant aux doux noms d'*Ajax*, *Soap*, *REST*, *RSS*, *Perl*, *Python*, *PHP*, *Ruby* ou *GreaseMonkey*. L'objectif de ce document n'est pas de développer ce volet technologique. Nous en évoquerons seulement les grandes caractéristiques pour mieux comprendre les enjeux. Le Web 2 repose notamment sur les interfaces utilisateurs dites "riches" qui permettent des possibilités d'affichage, de configuration et de maniement plus puissantes que les pages html classiques du "Web 1", que ce soit pour le développeur ou l'utilisateur final.

Tim Berners Lee, co-inventeur du web, remarque : *"en fait, ce Web 2.0 utilise les standards qu'ont produit tous ces gens qui travaillent sur le Web 1.0 : le Document object model, HTML et SVG (Scalable Vector Graphics), HTTP... Le Web 2.0 construit des services en s'appuyant sur les standards du web, ainsi que sur JavaScript. (...)"*

Pour se structurer en réseau de bases de données et services collaboratifs, le web 2.0 utilise donc des technologies matures, âgées de 5 à 10 ans pour la plupart. L'évolution n'est pas tant dans la technologie elle-même que dans la façon de mixer ces technologies pour apporter des services et une ergonomie nouvelle à l'utilisateur.

Ainsi, l'**Ajax** (voir glossaire), que l'on cite souvent à propos des technologies phares du Web 2 n'est pas un nouveau langage mais un nouveau concept qui utilise les standards XHTML et CSS, DOM (Document Object Model) pour l'affichage dynamique, XML et XSLT pour l'échange de données ainsi que la récupération de données asynchrones via XMLHttpRequest. JavaScript intègre et combine tout ces langages. Le terme Ajax a été apporté par Jesse James Garrett, de l'agence web *Adaptive Path*.

L'Ajax permet notamment de ne recharger que certains éléments d'une page plutôt que la page entière. Lors de modifications sur une page, seuls les nouveaux éléments se rafraîchissent, économisant ainsi des ressources, du temps et de la bande passante. L'un des précurseurs grand public dans l'utilisation de cette technologie fut A9, le moteur de recherche d'Amazon : il n'était pas nécessaire de recharger la page à chaque fois que l'on ajoutait de nouvelles sources et moteurs à interroger. ¹⁵

Greasemonkey est également basé sur des scripts. Cette extension pour le navigateur *Mozilla Firefox* va encore plus loin dans le "pouvoir" accordé à l'internaute : en exécutant des scripts, ce plug-in ¹⁶ permet de modifier l'apparence des pages web visitées, d'y ajouter des fonctions (boutons, case à cocher, suppression des publicités...). L'utilisateur avancé peut même écrire ses propres scripts ou télécharger des scripts mis à la disposition de tous.

Le Web 2 est en fait une sorte de "web lego" qui permet d'innover via l'assemblage de briques : les technologies de base sont multiples, il est possible de créer de nouvelles applications à valeur ajoutée en les assemblant différemment ou plus rationnellement.

Les "mash-up", un mixage de plusieurs applications (voir glossaire) en est une des illustrations les plus visibles : avec les mashups, on agrège le contenu émanant de plusieurs sites, afin de créer un site nouveau. On utilise alors souvent l'objet XML, l'AJAX du côté de l'internaute, les RSS, et les API (Web Services) des sites mixés. L'API n'est donc pas seulement un dieu égyptien. Pour faire simple, on peut dire que c'est une interface de programmation mise à disposition des internautes (avertis) pour développer des applications. Une application pourra alors accéder à des programmes système pour envoyer, recevoir ou extraire des données. Ainsi Google publie l'API de *Google Maps* qui est utilisé par de nombreux développeurs pour créer des applications dérivées du service de cartographie de Google : par exemple la carte de suivi du télescope spatial *Hubble*, la météo dans le monde. ¹⁷

Le rôle du RSS (voir glossaire) est par ailleurs prépondérant. Ce format qui permet d'exporter et lire des données d'actualités au format XML se retrouve dans quasiment toutes les applications dites Web 2 et peut en constituer l'emblème sinon la caractéristique majeure avec les tags. Il permet de surveiller les flux de nouvelles informations issues de sites d'actualités, de blogs, de moteurs de recherche... bref de tout site web qui propose un contenu qui se renouvelle. Le RSS va aussi servir à exporter des *bookmarks sociaux* que sont les tags ou les favoris par exemple.

Si le Web 2 n'est pas une révolution technologique en soit, la relative simplicité de mise en place de ces techniques et leurs coûts de programmation réduits expliquent le très fort développement que connaissent toutes ces applications. Par ailleurs, la baisse constante du coût d'hébergement et de bande passante a facilité la diffusion de ce web nouvelle génération.

Mais il subsiste des effets de seuil. Des applications gigantesques comme *Flickr* de *Yahoo!*, offrant l'hébergement et le partage de photos, impliquent la gestion de vastes bases de données et des infrastructures de stockage importantes. Seules de grandes firmes comme *Yahoo! Inc.*, *Google*, *Microsoft* ou *AOL* par exemple, peuvent se départir efficacement de ce type de contraintes.

Les familles d'applications du Web 2.0

TYPLOGIE

Les familles d'applications Web 2.0 sont nombreuses et diverses, de nouveaux services apparaissent quotidiennement (sans garantie de pérennité...et bien souvent de façon redondante). Sans prétendre dresser une liste exhaustive des nombreuses possibilités, on peut dégager une typologie des applications selon différents domaines d'activité. Les applications spécifiquement dédiées à la veille et à la recherche d'information seront développées dans la partie suivante.

▪ Bureautique en ligne

Courant 2006, Google a lancé successivement plusieurs logiciels collaboratifs en ligne : un outil de gestion d'agendas (*Google Calendar*¹⁸), un tableur (*Google Spreadsheet*) et enfin, après le rachat de la société *Writely*, un traitement de texte (*Google Docs*)¹⁹. Ces logiciels permettent de créer, modifier, partager et publier en ligne des documents sur un espace de stockage. Depuis plusieurs mois, on assiste à une véritable déferlante de logiciels de bureautique en ligne, désignés parfois sous le terme de "web office" : *RallyPoint*, *AjaxWrite* (pour *Firefox* uniquement), *ZohoWriter*, *Thinkfree*... Outre la gratuité, l'intérêt de ces logiciels concurrents de la suite *Microsoft Office* est leur disponibilité sur toute machine connectée au réseau et tout système d'exploitation. Attention toutefois aux questions de sauvegarde, de compatibilité, de synchronisation des données et enfin de confidentialité...

▪ Emploi

Les blogs constituent un bon moyen pour promouvoir son CV, mettre en avant expérience et compétences et faciliter les échanges. Des "jobblogs" verticaux spécialisés par secteur ou région connaissent un grand succès (voir par exemple ceux du site et moteur *Moovement*²⁰). Le networking professionnel est également à la mode, à travers des sites comme *LinkedIn* (créé en 2003 aux Etats-Unis), *Viadeo* (ex *Viaduc*, d'origine française) ou *Xing* (ex *OpenBC*, à vocation européenne). Le principe est de tirer parti des réseaux de connexions sociales virtuelles afin

d'identifier des salariés ou employeurs potentiels, mais aussi d'étoffer son réseau de clients, partenaires, fournisseurs...

L'e-cooptation (*Cooptin*, *Jobmeeters*²¹) met le réseau au cœur de la démarche de recrutement, via un système de prime de cooptation, et peut aider les demandeurs à identifier les offres du marché caché.

▪ Commerce en ligne

Le client est roi et le Web 2.0 permet de rehausser son niveau de participation sur les sites marchands. On peut ainsi *tagger* ou annoter des références d'ouvrages dans le catalogue *Amazon*. Grâce à des sites comme *Zlio* ou la *Shoposphère* de *Yahoo!*²², les internautes peuvent recommander à leur entourage des produits et services...en touchant une commission sur les ventes réalisées. Des guides d'achat et comparateurs de prix comme *Ciao!* intègrent les avis de consommateurs et d'"experts". Le site *Wantuno* (inspiré de l'américain *Woot*) innove dans le domaine de la vente événementielle en vendant un seul produit nouveau high tech par jour, annoncé en podcast vidéo à des internautes prévenus par flux RSS²³. Certains analystes voient l'avenir du commerce électronique dans le concept de "CtoB" (Consumer to Business) inverse du BtoC...

▪ Géolocalisation

Google Maps a été l'un des premiers services de géolocalisation, en s'appuyant sur l'outil de recherche pour localiser sur des cartes des services de proximité (commerces, station service...). La version 2 du site *Géoportail* de l'IGN, disponible depuis fin mai 2007, s'enrichit de nouveaux services pratiques et citoyens (accès au cadastre, visualisation des zones à risque...). *Google Earth* fait l'objet de partenariats avec la *NASA* ou le programme des Nations Unies pour l'environnement. La géolocalisation peut s'adapter également au partage de photos (*Yahoo! Maps* et *Flickr*, ou le projet *Pikeo* chez *Orange*)²⁴.

▪ Recherche d'informations

Des agrégateurs de média comme *Digg* ou *Wikio* laissent une large place à l'internaute pour *tagger* ou recommander par vote les articles issus de médias classiques ou blogs. On suivra également l'actualité en syndiquant des flux RSS à travers des agrégateurs comme *Netvibes* ou *Newsgator*. Les grands moteurs (*Google Custom Search*, *Yahoo Search Builder*, *MSN Live search Macros*) offrent des options de

création de moteurs personnalisables, effectuant des recherches sur des bouquets de sources présélectionnées par l'utilisateur. On peut mutualiser les connaissances à travers des wikis collaboratifs généralistes (*Wikipedia*) ou spécialisés. Enfin, le projet de moteur collaboratif *Wikiasari*, qui doit aboutir courant 2007, aura recours aux mêmes technologies libres que la célèbre encyclopédie, ainsi qu'à un réseau humain ou "communauté de confiance" qui validera les résultats de la recherche.

The screenshot shows the Wikio website interface. At the top, there is a search bar with the text 'rechercher' and a dropdown menu for 'Aéronautique' and 'Toute l'information en français'. Below the search bar, there are navigation tabs for 'Actu', 'Blogs', and 'Shopping'. The main content area displays a list of news articles under the heading 'Aéronautique'. The first article is 'Airbus s'engage à aider ses fournisseurs', followed by 'Aéronautique: L'A380 a atterri à Roissy', 'Des pièces détachées de Concorde bientôt mises aux enchères', 'Premier atterrissage de l'Airbus A380 à Roissy', and 'Conférence sur les enjeux stratégiques et perspectives de l'aéronautique mondiale - Jeudi 15 novembre 2007 - Paris'. On the right side, there is a sidebar with a 'publier un article en direct' button, a 'tags' section with a 'liste complète' link, and a list of various categories and tags including 'France', 'International', 'Politique', 'Economie', 'Finance', 'Santé', 'Société', 'High-tech', 'Science', 'Culture', 'Loisirs', 'People', 'Sport', and 'Podcast'. There is also a 'proposer votre site' section with options to 'ajouter les news sur votre site' and 'ajouter un lien sur votre site'.

▲ Les actualités sur Wikio : sites de presse, blogs et navigation par tags (à droite)

• Diffusion d'actualités

Pour Pierre Chappaz, fondateur du site *Wikio*, "les blogueurs ne sont pas journalistes, mais ils peuvent apporter une vraie expertise sur certains sujets qu'ils connaissent, et il serait dommage de s'en priver". Selon lui, les médias 2.0 ne seront plus des canaux exclusifs de diffusion mais producteurs de contenu et des "marques". Désormais, le scoop n'est plus l'apanage des seuls journalistes : aux Etats-Unis, une jurisprudence donne droit depuis 2006 aux blogueurs de protéger leurs sources, au même titre que les journalistes professionnels. Des centaines de blogs politiques ont fleuri sur le web, certains d'entre eux se révélant particulièrement influents. On parlera de journalisme citoyen à propos de sites comme *Agoravox* ou de l'"allée numérique" *Rue 89*, lancée en mai 2007 par d'anciens journalistes du quotidien *Libération*. Ces nouveaux médias en ligne marient les communications venant de journalistes professionnels, "experts" et internautes.

▪ Loisirs

On retrouve ici la notion de "filtrage collaboratif", basé sur un système de recommandations et d'évaluations par les utilisateurs. Ces recommandations peuvent concerner les livres, les destinations de voyage, la musique, la vidéo...et proviennent de l'analyse de comportements d'achats des internautes et de leurs évaluations de produits. Le moteur "d'analyse prédictive" *Criteo*²⁵ a ainsi lancé en février 2007 un service de recommandations automatique de blogs, après avoir utilisé sa technologie pour l'évaluation de vidéos.

La participation des internautes peut aller jusqu'au financement et à la production d'œuvres artistiques : le projet "A *swarm of angels*" permet d'aider à la création de films qui seront distribués sous licence *creative commons* et diffusés via les réseaux peer-to-peer.

▪ Partage de documents multimédia

Il s'agit sans doute d'une des applications les plus répandues du Web 2.0. Les grands acteurs du web ont largement investi ce domaine : Google a racheté en 2006 le site de partage de vidéo *Youtube*, *Yahoo!* de son côté avait racheté le site de partage de photos *Flickr*. Les favoris peuvent également être mutualisés via des services en ligne comme *Del.icio.us* (groupe *Yahoo!*) ou *Blogmarks*. De nombreux services de partages de photos, vidéo ou podcasts se sont développés récemment : on peut citer parmi eux les sites français *DailyMotion* ou la banque d'images *Fotolia*. *Podemus* ou *Radioblog* figurent parmi les sites communautaires de podcasts.

▪ Enseignement et recherche

Henry Jenkins²⁶, professeur au MIT, imagine des unités d'enseignement conçues sur les modèles *YouTube* ou *Wikipedia* : des départements qui "permettraient le déploiement rapide d'expertises dispersées et la reconfiguration des champs". De nouveaux usages apparaissent dans le paysage éducatif, mais attention là encore à l'illusion de facilité : selon le philosophe Fernando Savater, « *Internet n'éduque pas, il nécessite même une éducation préalable* ». Olivier Ertzcheid, dans son blog *Affordance*²⁷, distingue plusieurs éléments caractéristiques d'une "science 2.0", et notamment la possibilité d'annoter et commenter des publications scientifiques (notamment les collections *Open Access Plos One*²⁸) ou de nouveaux modes d'interrogation des bases de données (interface *Hubmed* sur la base médicale *Medline*).

Marie France Blanquet²⁹, maître de conférence en sciences de l'information, distingue le travail coopératif du travail collaboratif. En matière d'apprentissage, le travail coopératif réalisé par chaque équipe contribue à une œuvre collective, où le rôle de chacun est défini de façon claire. L'apprentissage collaboratif résulte du travail individuel soutenu par des activités de groupe ou d'équipes, de façon souple et ouverte. Ce type de travail se base sur les capacités de communication et d'interaction de chacun. L'écriture collective des articles composant l'encyclopédie Wikipédia en est un exemple.

Plusieurs expérimentations d'usage des blogs et podcasts ont été tentées à l'Université. A Lyon II, des étudiants ont pu enregistrer des cours sur un Ipod, pour diffusion en podcast sur leurs blogs. A l'Université Catholique de Lille, les étudiants de Jean-Paul Pinte³⁰, chargé de l'innovation pédagogique, doivent alimenter régulièrement un blog. Pour ce chercheur spécialiste de la veille en éducation, le renouveau des méthodes d'enseignement se traduit par le rôles des **"Enseignants Connecteurs et étudiants pronétaires"**.

▪ Bibliothèques

La "bibliothèque 2.0" place elle aussi l'utilisateur au centre, et ce à travers plusieurs types d'expérimentations : de l'utilisation des logiciels de messagerie instantanée pour échanger avec les utilisateurs, à l'intégration de tags pour enrichir les catalogues, en passant par des interfaces de consultation simplifiées... De nombreux blogs de bibliothécaires (*Bibliobsession*³¹) permettent de suivre l'état de l'art. La bibliothèque de l'Université d'Albany est allée jusqu'à publier un manifeste du bibliothécaire 2.0³².

COMMENT IDENTIFIER UN SITE WEB 2.0 ?

Est-ce que le site sur lequel vous naviguez ou l'application que vous utilisez est "Web 2.0" ? Quelle importance pourrait-on répondre...Toutefois, un site de cette génération offrira plus de possibilités à vous, l'utilisateur ! Pour reconnaître un tel site, quelques caractéristiques sont à retenir :

- le type de site : il appartiendra à l'une des grandes familles citées plus haut : il sera blog, bookmark ou réseau social, *mashup* ou page personnalisable : bref, il permet de partager et de diffuser du contenu aisément. Le site peut être lui-même une application Web 2 ou en intégrer des "briques". Exemple: un site Corporate qui héberge des blogs de ses salariés et/ou diffuse son actualité via un fils RSS.

- le rôle de l'utilisateur : le site Web 2 vous permet de participer, de personnaliser et de collaborer : vous pouvez en effet ajouter des commentaires, écrire un billet, lire un flux RSS, noter (voter) ou annoter un article, télécharger ou "uploader" des images, des vidéos, écouter un podcast. Ces sites proposent souvent des interfaces personnalisables : couleurs, style, organisation de l'affichage du contenu....

- la navigation : elle est plutôt simplifiée par rapport aux sites de première génération. Les menus sont plus légers, un flux RSS est toujours proposé et un nuage de tags est souvent présent.

- le testeur c'est vous : non content de faire apporter une partie de son contenu par l'internaute, un service Web 2.0 fait en plus souvent tester ses applications à ses utilisateurs et est donc fréquemment en version bêta (voire alpha): une bêta publique ou privée (l'internaute fait alors partie d'un petit cercle d'initiés, dans le secret des dieux).

Les applications Web 2.0 pour la veille et la recherche d'informations

LES TYPES D'APPLICATIONS

Pour le veilleur et le chercheur d'information, les applications du Web 2.0 sont riches, tant dans les possibilités offertes par leurs interfaces que dans le contenu diffusé. Mieux connaître ces services permettra de trouver davantage d'information (ou autrement), de gagner du temps, de mieux partager voire d'être plus visible sur le web.

Vous l'aurez compris, le web 2.0 se fonde essentiellement sur la notion de partage, de participation et de collaboration entre les individus. Ce n'est pas forcément toujours facile, possible voire autorisé dans le monde professionnel. Mais partager c'est aussi à terme pouvoir recevoir...

I. Partager et collaborer

a. Les Bookmarks sociaux

Les bookmarks sociaux sont une des applications les plus caractéristiques du Web 2 et en sont précurseurs. Apparus en 2004, ces services vous offrent la possibilité de partager vos bookmarks c'est-à-dire vos favoris. Ainsi, après enregistrement le plus souvent gratuit, vous pouvez mettre en ligne tout ou partie de vos favoris (titre,

adresse et description d'une page ou site) et les rendre potentiellement accessibles aux internautes du monde entier, et du moins, aux utilisateurs de ce service. Vous avez la possibilité de garder ces favoris privés en ne les partageant qu'avec vos proches ou seulement...vous-même.

Ces applications sont au départ très simples et proposent ensuite, suivant les sites, plus ou moins de perfectionnements : il suffit d'ajouter le nom et l'adresse (url) de la page web qui vous intéresse, assortis des commentaires de votre choix. Ces adresses web sont rattachées à un espace qui vous est personnel, auquel vous accédez en vous identifiant. Cet espace personnel est identifiable via une adresse précise. Par exemple, sur *Del.icio.us*³³ (de Yahoo!), l'un des services de bookmarks sociaux les plus utilisés, l'url désignant mes favoris est <http://del.icio.us/slide68> . En naviguant sur *Del.icio.us*, vous pouvez accéder aux favoris mis en ligne par des collaborateurs, amis ou inconnus qui partagent les mêmes intérêts et passions que vous. Leurs pages sont également identifiées par une url unique.

L'ajout d'adresse

L'ajout d'adresse s'effectue en se connectant directement sur le site de bookmarking social ou, plus pratique, via une barre ou un bouton dédiés qui, intégrés à votre navigateur facilitent l'ajout du site que vous êtes en train de visiter. Vous pouvez également trouver des icônes sur des pages web vous proposant d'ajouter un contenu, un site, à votre service de social bookmarking. Peut-être avez-vous déjà rencontré ces icônes à la fin d'un billet de blog par exemple (voir illustration ci-dessous). Ils permettent alors d'ajouter l'adresse du billet à un ou plusieurs services de bookmarks sociaux ainsi qu'à des services d'actualités collaboratives (nous en reparlerons plus bas).

Posté à 03:23 le 7/01/2007 dans [Varia](#)

Pour vous souhaiter une bonne année 2007, Blog2B a doublé tous les quotas d'espace disque. Les blogs gratuits disposent donc de 200 Mo, ceux à 5 euros de 500 Mo et ceux à 10 euros de 2 Go! De quoi laisser libre court à votre créativité!

Pensez tout de même à compresser un minimum vos images et vidéos, sans quoi l'espace risque de disparaître rapidement (et vos lecteurs vont se lasser de patienter pour le chargement!). La méthode la plus simple pour diminuer la taille de vos images? Installez [ImageResizer](#), un plugin gratuit offert par Microsoft et qui permet de réduire la taille des photos d'un simple clic droit.

[Commentaires \(2\)](#) | [Poster un Commentaire!](#) | 

▲ Ce blog propose, via des icônes, l'ajout de ce billet dans les favoris de Del.icio.us ou les services sociaux d'actualités.

Lorsque vous ajoutez une adresse, vous lui attribuez des *tags* (voir glossaire) c'est-à-dire un ou plusieurs mots clés de votre choix qui vont caractériser, expliquer, contextualiser, étiqueter le favori et donc l'identifier de manière simple pour vous et tous les visiteurs. C'est une information qui vient s'ajouter à la traditionnelle description du site, C'est donc vous qui choisissez votre propre mode de classement, de hiérarchisation et de labellisation (et non plus un service qui vous propose de ranger vos ressources dans des catégories prédéfinies). Cette possibilité fait partie de ce que l'on appelle la folksonomie, à savoir une taxonomie (science du classement) "populaire" en quelque sorte.

Ainsi, vous pouvez ajouter l'adresse du site portail *NanotechNow*, <http://www.nanotech-now.com/>, consacré aux nanotechnologies et y adjoindre les tags "nano", "nanotehnologies", "sciences", "science", "research" mais aussi pourquoi pas "qualité", "excellent" pour donner votre opinion.

The screenshot shows a Del.icio.us profile page for user 'phaze9' with the tag 'china'. The page header includes the user's name, navigation links, and a search bar. The main content area displays a list of bookmarks, each with a title, a snippet, and a 'save this' link. The sidebar on the right contains two sections: 'related tags' and 'capital', both listing various tags with their respective counts.

▲ Des favoris consacrés à la Chine dans *Del.icio.us*. A droite, les tags associés à ces bookmarks.

Les tags

Sur une application typique de social bookmarking (type *del.icio.us*), il est possible de rechercher sur le titre des favoris de tous les autres utilisateurs et **aussi sur leurs tags**. Ainsi, en cliquant sur le tag "*Nanotech*", vous obtiendrez la liste de tous les favoris que les internautes ont classés sous ce label. C'est là l'un des principaux intérêts : la possibilité de rechercher sur le nom des étiquettes textuelles définies par les internautes eux-mêmes. Mais, tous les internautes n'ont pas la même logique de classement et il est fort possible de passer à coté de favoris intéressants mais étiquetés selon un mot auquel vous n'auriez jamais pensé. De plus, un même mot peut avoir plusieurs sens ou être employé différemment selon les individus. *Del.icio.us* a résolu –partiellement- ce problème en offrant la possibilité de décrire ses tags, c'est à dire de les expliquer (par *orange*, je parle de la couleur ou de la production d'agrumes en Floride). C'est donc un début de sémantique appliqué à cette folksonomie. Quoiqu'il en soit, le problème du sens des mots employés par les créateurs de contenu concerne l'ensemble du web : Lorsque vous ne trouvez pas un site via une requête sur un outil de recherche et que ce site existe pourtant bel et bien, c'est souvent parce que vous n'employez pas les mêmes mots clés que ceux utilisés par le créateur du site dans son texte.

Les tags peuvent également s'appliquer à des photos (*Flickr*), à des projets (*43Things*) ou des vidéos (*YouTube*, *DailyMotion*). Ils peuvent servir aussi à étiqueter et catégoriser les billets d'un blog.

Les communautés

L'autre atout des bookmarks sociaux : pouvoir **constituer et se connecter avec une communauté de spécialistes**. Par *spécialiste*, on désignera simplement un internaute qui centre ses recherches sur les mêmes sujets que vous, qu'il soit amateur ou professionnel.

Exemple : Votre domaine, ce sont les biotechnologies. En entrant la requête "Biotechnologies" sur le moteur de recherche de *Del.icio.us*, vous obtenez une liste de plusieurs milliers de favoris. Pour chacun d'entre eux, est indiqué le nombre de personnes qui ont sauvegardé cette même adresse. En sélectionnant l'un de ces utilisateurs, vous pouvez découvrir un internaute qui, comme vous, régulièrement, ajoute des adresses de sites dédiés ou connexes aux *biotech*. Il ne vous reste plus qu'à mettre en surveillance ce membre de *del.icio.us*. Il disposera en effet de sa propre adresse XML. Exemple : <http://del.icio.us/rss/vantive98>. Vous serez alors averti dès lors que cet utilisateur ajoute un favori à sa page.

La valeur des outils de social bookmarking est encore plus probante lorsqu'ils sont dédiés à des communautés de pratiques ou à des secteurs d'activités bien précis. Plusieurs initiatives ont déjà été mises en ligne³⁴. On peut citer *Connotea*, édité par la célèbre revue *Nature* et donc destiné à un public scientifique. Lors de la mise en favoris, ce service extrait automatiquement les références bibliographiques lorsqu'elles sont issues de sites tels que *Nature*, *Science* ou *PubMed*. Plus ancien, *CiteUlike* est basé sur le même principe mais est "compatible" avec davantage de revues scientifiques encore. *Snipitron* est lui dédié aux chercheurs, étudiants et professionnels. Ici, vos bookmarks font partie d'un projet qui s'écoule dans le temps, concernant une recherche ou une tâche précise.

L'intérêt des bookmarks sociaux est donc d'abord dans l'utilisation de la "communauté" afin de trouver des sources pertinentes. Ce type d'outil révèle toute sa puissance lorsqu'un utilisateur proche de vous en terme de domaines de veille a été identifié. Vous pouvez donc alors l'utiliser comme "collaborateur virtuel" mais aussi rechercher parmi les sources jugées pertinentes par un groupe d'utilisateurs, identifier des leaders de veille ou juger de la popularité d'une source (combien de fois a-t-elle été mise en favoris ? Ce nombre de sauvegardes n'étant qu'un indicateur parmi d'autres). Si vous ne partagez pas, ces outils restent toutefois un très bon moyen de mémoriser et catégoriser vos favoris en un point unique d'accès.



Quel service ?

Il existe plusieurs dizaines de services de bookmarks sociaux. Ils proposent des fonctionnalités plus ou moins sophistiquées. Un des critères essentiels de choix est leur succès. En effet, leur force repose en grande partie sur le nombre de favoris d'utilisateurs partagés, à moins de fréquenter des bookmarks dédiés à des communautés très ciblées. D'autre part, la rapidité du service et la pertinence de son moteur de recherche sont aussi des paramètres importants. *Del.icio.us*, qui appartient à *Yahoo!* est l'un des plus connus et des plus utilisés. *Yahoo! MyWeb2* rassemble également une grande communauté ainsi que *Furl* ou encore *Blogmarks* et *Blinklist*³⁵. Certains de ces sites, au-delà de l'archivage de l'adresse, proposent de conserver une copie intégrale d'une page web. D'autres vous suggèrent des sites similaires à ceux que vous avez mis dans vos favoris. Cette similarité se base notamment sur les favoris ajoutés par les autres utilisateurs. La recherche par similarité (ou par suggestion, en fonction de vos habitudes de navigation) offerte par des services comme *Yoono*, *StumbleUpon* ou *Similicious*³⁶ est une méthodologie de recherche particulièrement intéressante. Nous en reparlerons plus loin.

Pour choisir précisément votre service de bookmarks sociaux, vous pouvez vous aider de ces **ressources**³⁷ :

- le site *Irox* propose un tableau comparatif de près de 20 services
- le blog de *Robin Good* offre une sélection intéressante de bookmarks sociaux
- la page *Free Bookmark Managers* affiche une liste d'une quarantaine de logiciels et utilitaires dédiés au partage et à la gestion de vos favoris (bookmarks) en ligne via accès distant, en local ou sur serveur. Attention, tous ne sont pas sociaux, c'est-à-dire basés sur le partage.



b. Le Multimédia : photos et vidéos

Le partage de vos ressources ne se limite pas au texte et adresse de sites. Il est également possible de partager vos photos et vidéos.

Les images

*Flickr*³⁸, racheté par *Yahoo! Inc.* en 2005, est l'application type de partage d'images voire même l'archétype du service Web 2.0 : ce site permet de mettre en ligne ses photos, de les partager avec ses proches, ses collègues ou de les laisser en accès privé.

Il est possible d'utiliser un moteur pour rechercher à travers les millions de photos mises en ligne par les autres utilisateurs.

Flickr n'a pas de points communs avec les nombreuses applications Web 2 en bêta à l'interface artisanale. Sur le service de *Yahoo!*, la puissante interface en fait un outil très simple à utiliser. Ainsi, via le menu *Organize*, vous pouvez accéder à une véritable table de montage virtuelle pour organiser vos photos.

The screenshot shows a Flickr search interface. At the top, there are tabs for 'Photos', 'Groups', and 'People'. A search bar contains the text 'exxon' and a 'SEARCH' button. Below the search bar, it says 'We found 1,422 photos matching **exxon**.' There are three search results visible:

- Result 1:** A graphic with the text 'EXPOSE EXXON'. The user is 'dilligaf?' and the photo was uploaded on 29 July 2005. The tag 'exxon' is associated with it.
- Result 2:** A photograph of an Exxon gas station sign. The user is 'dmuth' and the photo was uploaded on 27 April 2007. The tags 'katrina', 'greed', and 'exxon' are associated with it.
- Result 3:** A map titled 'Exxon Valdez Comparison Map, pay up.' The user is 'palustrine' and the photo was uploaded on 15 October 2006. The tag 'exxon' is associated with it.

On the left side of the search results, there is a 'Y! SEARCH' widget with a search bar containing 'exxon' and a 'Search the Web' button. Below it, it suggests other search terms: 'Also try: gas, oil, mobil, war or esso'.

▲ Une recherche sur Flickr. A droite, les noms d'utilisateurs et les tags associés aux photos.

Comme sur les bookmarks sociaux, la recherche peut s'effectuer sur le titre ou la description mais aussi sur les tags : chaque internaute qui partage ses photos leur attribue des mots clés pour les définir, les expliquer, les contextualiser voire donner son opinion.

Chaque liste de résultats, chaque photo est accompagnée par ses tags. Ceux-ci permettent aussi d'affiner ou d'élargir votre recherche lorsque vous en sélectionnez un. Il est possible de rechercher sur les photos, les utilisateurs mais aussi sur les groupes. En effet, la notion de communauté est également importante. Les "groupes" rassemblent les utilisateurs qui partagent les mêmes centres d'intérêt photographiques. Des forums de discussions sont intégrés à ces groupes, permettant d'entrer en contact direct avec certains membres.

Par exemple le groupe "The Eyes Have It" ³⁹ rassemble les utilisateurs (et donc leurs photos) autour du thème de la communication visuelle des secteurs pharmacie et biotech.

Que peut trouver le veilleur sur un service de photos tel que *Flickr* ?

Finale­ment, la même chose que sur les bookmarks sociaux : tout ce que les membres ont voulu partager : des photos du dernier salon automobile de *Detroit* ⁴⁰ ou de la convention annuelle des industries cosmétiques d'Amérique du Nord. La veille concurrentielle ou la veille image sur votre propre société est également intéressante : Ainsi, une recherche (*full text*) sur la société pétrolière *Exxon* affiche plus de 1400 photos relatives aux stations services, aux puits de pétrole, aux raffineries, aux employés, aux manifestations de contestataires ou aux publicités détournées.

Comme avec le social bookmarking, vous pouvez vous attacher à "suivre" les membres qui partagent les mêmes centres d'intérêt que vous ou votre organisation : après sélection de la page de l'utilisateur, la mise en surveillance de son flux RSS permet d'être averti dès que cette personne ajoute de nouvelles photos.

D'autres services permettant le partage de photos existent. Toutefois, ils ne vont pas aussi loin que *Flickr* dans la logique Web 2.0. Le service **PhotoBucket** ⁴¹ est le grand concurrent de *Flickr* mais avec une approche plus grand public. *Flickr* est davantage fréquenté par des internautes avertis, des professionnels et des passionnés de photographie. En ce sens, le service de Yahoo! sera donc plus intéressant pour le veilleur.

Squidoo ⁴² est en quelque sorte un annuaire de bookmarks perfectionné, une rencontre de services sociaux comme *Del.icio.us* et *Flickr*. *Squidoo* vous permet en effet de créer une page de liens dans un domaine qui vous passionne et de la structurer en rubriques agrémentées de photos.

L'interface, simple, rend possible la création rapide de pages spécialisées (appelées "lens">objectifs, lentilles) et de les rendre visible à tout un chacun. On trouve par exemple plusieurs "lens" consacrées aux **nanotechnologies** ⁴³. Chaque *lens* est accompagnée de tags connexes à son sujet.

Vous pouvez rechercher ces pages grâce à un moteur de recherche, un répertoire par catégories, un nuage de tags (exemple : le tag *Marketing* ⁴⁴) ou même parcourir le *Top100*.

La vidéo

Via les services *YouTube* ou *DailyMotion*, le partage de vidéos en ligne est plus médiatisé que les services de photos. Notamment parce que le leader Google a racheté *YouTube* en octobre 2006 générant un buzz sans précédent au sein des blogs et de la presse "traditionnelle" (Mainstream medias).

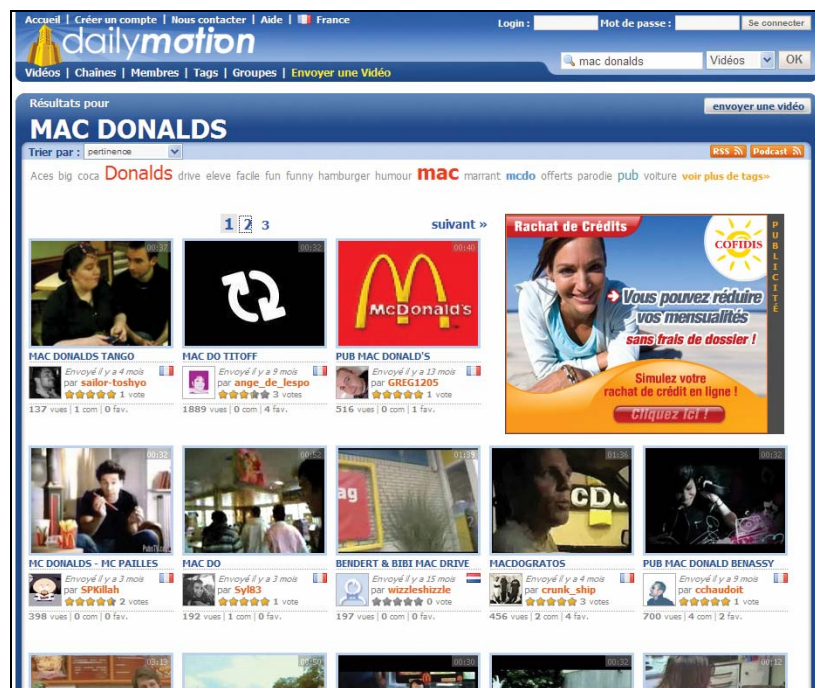
*YouTube*⁴⁵ et ses concurrents Web 2.0 fonctionnent à peu près de manière équivalente à *Flickr* sans en atteindre toutefois sa clarté ergonomique.

Tout ces sites de vidéos vont permettre d'"uploader" (charger vers le site) et partager vos vidéos (converties automatiquement au format *FlashVideo*), et aussi bien sûr de rechercher parmi les vidéos des autres utilisateurs.

Sur ces applications, on retrouve de nombreuses caractéristiques du Web 2 : les vidéos, classées par catégories, sont accompagnées de *tags* qui aident à les contextualiser et les décrire, en plus du titre et de la description renseignés par l'utilisateur. Un fils RSS permet de surveiller l'ajout de vidéos dans une catégorie ou par un utilisateur défini. A l'instar des billets d'un blog, les internautes peuvent commenter les vidéos. De plus, les vidéos sont notées, ce qui impacte ensuite leur place dans le classement et donc leur visibilité.

Ces services contiennent des favoris intégrés où l'on peut sauvegarder ses vidéos préférées. Il est possible de partager sa vidéo par mail ou messagerie instantanée, de la poster sur son blog. D'ailleurs, de nombreux blogueurs utilisent les services de type *YouTube* (ou *Flickr* pour la photo) pour stocker leurs fichiers multimédias, trop volumineux pour les hébergeurs de blogs. La notion de groupe, comme sur *Flickr* est présente : des communautés rassemblent des utilisateurs autour de grandes thématiques liées par exemple à la politique, aux sports et aux voyages.

YouTube a de nombreux concurrents, aux fonctions très proches. Impossible de tous les mentionner. Mais on peut retenir cependant le français *DailyMotion*⁴⁶, très populaire, d'ailleurs né quelques semaines avant *YouTube*. Citons également *Yahoo! Videos*, *AOL Videos*, *Rever*, *Grouper* (de Sony), *Metacafe*, *SoapBox* (de Microsoft), *Zudeo*, *Wideo*, *MySpace*, *Vimeo*⁴⁷... Certains de ces services proposent du contenu amateur ainsi que des films professionnels, issus des networks de TV et de cinéma. Mais les plus petits services se cantonnent aux vidéos envoyées par l'internaute, que celles-ci aient d'ailleurs été copiées ou non.



▲ Une recherche sur *DailyMotion*.

Quel contenu à usage "professionnel" le veilleur peut espérer trouver sur ces plateformes? Essentiellement des vidéos consacrées à des sociétés et individus. Ainsi, si je réitère ma recherche sur la société pétrolière *Exxon*, j'obtiens des vidéos sur des manifestations, des marées noires, le réchauffement climatique, des actualités boursières, des retransmissions de procès, des détournements de publicité....La recherche de vidéos s'averra particulièrement utile pour une veille Image, Produits ou concurrentielle. Le contenu de vidéos est également souvent très politique : à partir de la campagne des élections présidentielles françaises de 2007, la diffusion de vidéos sur les hommes et femmes politiques est devenue une véritable arme. Autre aspect intéressant, la possibilité de découvrir de nombreux extraits de vidéos de documentaires, sur des sujets extrêmement variés : écologie, industrie, société, consommation...

Le contenu de ces services de vidéos est donc, à l'image du net, hétérogène mais potentiellement riche.

c. Partage de CV et réseaux sociaux



Terminé le temps où seules les personnes à qui vous transmettiez votre CV ou votre carte de visite pouvaient prendre connaissance de votre identité et de votre parcours professionnel. Via les plateformes de réseaux sociaux, le push fait place au pull, de plus dans une optique de cooptation et de recommandation : il y est possible de mettre son CV en ligne sous format hypertexte et de demander à rentrer en relation avec des proches de votre réseau immédiat. Les plateformes de réseaux sociaux sont nombreuses : certaines sont à vocation généraliste, d'autres uniquement dédiées aux actifs, d'autres encore réservées à une communauté (étudiants, lycéens, anciens élèves, passionnés de voyages).

Ces applications représentent un marché conséquent puisque pour les seuls USA, on prévoit un chiffre d'affaires des réseaux sociaux de 2,5 milliards de dollars à l'horizon 2011 contre 350 millions en 2006.⁴⁸ La France est dominée par la plateforme Viadeo (ex-Viaduc)⁴⁹. Ce service leader lancé en 2004 compte plus de 750 000 profils. Dans les autres pays, c'est LinkedIn⁵⁰ qui domine généralement avec plus de 9 millions d'utilisateurs dans 135 pays.

The screenshot shows the Viadeo search interface. At the top, there are navigation tabs: Accueil, Recherche, Espace Emploi, Mon profil, Mes contacts, Mises en relation, Hubs, and Annonces. Below these are sub-tabs for search filters: Standard, Anciens élèves, Rapide, Mémorisée, Profils consultés, Profils annotés, Blogs des Membres, Communautés, and Résultats. The main content area displays '57 membres correspondent à votre recherche' and includes a sidebar with suggestions and filters, and a main list of search results.

Suggestion

- Etudes Marché
- Etudes de marché
- Etudes de marché sectorielles
- Etudes de marché pays produits
- Marché public
- Part de marché
- Nouveau marché
- Le Bon Marché
- Marché
- Marché libre

Votre recherche

Mots-clés : Etudes de marché
Région : Ile-de-France
Pays : France
Secteur : Pharmacie et santé - Pharmacie

Mémoriser cette recherche

En cliquant sur **Mémoriser**, vous pouvez créer un agent de recherche pour relancer celle-ci régulièrement.

Réseau d'un membre

Il s'agit du nombre total de personnes avec qui ce membre peut entrer en relation en passant par l'intermédiaire d'autres membres.

Distance

Je désire recevoir chaque semaine, par mail, les nouveaux résultats de cette recherche.

Analyste **Etudes de Marché** Hôpital, MSD (Merck Sharp & Dohme - Chibret)
Ancien élève : Université Paris 5 René Descartes - Lycée Janson de Sailly
5 filleuls - Son réseau : 66 500 membres - Distance > 4 - Entrer en relation

Consultant en dossier de Prix et Remboursement et Accès au **Marché**, AREMIS Consultants
Ancien élève : Université Paris 6 Pierre et Marie Curie - Université Paris 5 René Descartes - Université Paris Dauphine - Université Rennes 1 - Lycée Assomption
1 filleul - Son réseau : 92 000 membres - Distance > 4 - Entrer en relation

Chef de produit, Pfizer - Chef de produit, Pfizer
1 filleul - Son réseau : 36 200 membres - Distance > 4 - Entrer en relation

Chef de produit, Pfizer - Chef de produit, Pfizer
1 filleul - Son réseau : 36 200 membres - Distance > 4 - Entrer en relation

▲ Une recherche sur Viadeo.

Les fonctionnalités

Que trouve-t-on sur ces sites ? Notamment des profils de personnes en activité ou en recherche d'emploi ainsi que des "hubs" regroupant des utilisateurs autour d'un même centre d'intérêt, toujours la notion de communautés de pratiques propre au Web 2.0. Sur ces services, tout est pensé pour créer du lien social. Le profil individuel se constitue d'un CV enrichi de nombreux liens vers ses employeurs, ses amis, ses réseaux professionnels et associatifs, ses travaux et publications. Les *hubs* regroupent les personnes autour de secteurs, métiers, associations d'anciens élèves, pays, régions. Par exemple, sur *Viadeo*, s'affichent les hubs "*Analyse, Veille Stratégique & Concurrentielle*", "*Intelligence Eco*", "*Blogs & Bloggers*", "*Editeurs de Logiciels*", "*Profession: documentaliste and co*" mais aussi "*Grande distribution*", "*Océanie*", ou "*Normes, méthodes et outils*". Dans ces hubs, des forums organisent les discussions des membres autour de thématiques définies. Un moteur de recherche permet de trouver des profils selon un secteur, une fonction, une région. Une fois un profil identifié, il faut demander à entrer en relation avec lui. Cela sera d'autant plus facile et rapide s'il est peu éloigné de votre propre réseau de contacts.

La veille

L'intérêt pour le professionnel de l'information est double : il peut augmenter sa visibilité et celle de certains de ses projets, services, publications, thèses ou articles. D'autre part, il peut rechercher des données sur des personnes et entreprises ou identifier un "expert". Comme sur d'autres services Web 2.0, il est facile de se constituer une communauté d'utilisateurs partageant ses centres d'intérêt, que l'on visitera régulièrement. Ainsi, sur *Viadeo*, une recherche sur le secteur *Agro-alimentaire* avec le mot clé *veilleur*¹ affiche plus de 50 résultats avec les fonctions, entreprises et synthèses de parcours des membres. Vous pouvez alors entrer en relation avec ces personnes ou tout simplement collecter des données sur celles-ci et leurs réseaux. Ces sites sont d'ailleurs devenus un outil essentiel pour certaines fonctions : commerciaux pour la prospection, recruteurs, journalistes et...responsables de casting pour émissions TV.

¹ Certaines fonctionnalités avancées sont payantes.

Les autres réseaux

Parmi les autres services de réseaux sociaux, outre *LinkedIn* davantage tourné à l'international, il existe, assez utilisés en France, *6nergies.net* et *Xing* (ex *Open BC*)⁵¹. Les réseaux ciblant plus particulièrement les jeunes (*FaceBook*, *MySpace*) peuvent également être utilisés. *MySpace* fédère plus particulièrement des professions de la communication, de la publicité et des métiers multimédias.

Le réseau *FaceBook*⁵² est original : ce site nord-américain est précurseur d'un social networking très segmenté, réservé aux étudiants et aux lycéens. Au départ, les membres devaient s'inscrire obligatoirement via l'email de leur université, afin de garantir l'authenticité des profils et donc le sérieux de la communauté. Depuis, *FaceBook* s'est ouvert aussi aux entreprises. Ainsi, ce site capte très tôt les futurs actifs et les fidélise.

Des ressources

Certaines ressources peuvent vous aider à identifier les plateformes de réseaux sociaux professionnels utiles pour vos recherches : Le wiki "*List of social networking websites*" liste de nombreuses plateformes de réseaux sociaux⁵³ et le site "*Link n Rank*" catégorise des sites de social networking⁵⁴.

d. Les sites d'actualités collaboratives : recommandation et notation

Ces sites recouvrent en fait différents types d'applications :

- Les "Digg Like"

Ces services, qui s'inspirent tous du précurseur *Digg*⁵⁵, recueillent des actualités proposées par les internautes. Le principe : après enregistrement, un formulaire simple vous permet de poster l'actualité de votre choix, susceptible d'intéresser les autres internautes : son titre, son url, sa description. Il faut aussi choisir une catégorie. Par exemple : *Technologies, Sciences, Monde, Affaires, Divertissement*. Ces catégories peuvent être subdivisées en thématiques plus fines. L'actualité, si elle est validée par les modérateurs, apparaîtra alors en ligne. Elle sera notée par les autres internautes qui votent pour elle selon son intérêt. Les actualités disposant du plus grand nombre de points se placent alors en haut du classement.

The screenshot shows the Digg website interface. At the top, there are navigation tabs for 'News', 'Videos', and 'Podcasts'. Below that, there are links for 'Join Digg', 'Login', and a search bar. The main content area is titled 'General Sciences' and features a list of articles. Each article entry includes a 'digg' count in a yellow box, the article title, a brief description, and a 'More...' link. The articles are ranked by their 'digg' counts: 'Genome of DNA Pioneer Is Deciphered' (86 diggs), 'Humans have a bit of shark in them' (402 diggs), 'Calculate pi by throwing frozen hot dogs' (484 diggs), and 'How to Heal Wounds Faster' (548 diggs). On the right side, there are two promotional boxes: 'What's Digg?' and 'Flash API Contest'.

▲ La catégorie Sciences sur Digg. A gauche, les nombres de votes.

Ces sites affichent les principales caractéristiques du Web 2.0 : interface épurée, possibilité de laisser des commentaires, fils RSS pour surveiller les nouvelles actualités, classement des articles les plus lus, les mieux notés, possibilité d'envoyer par mail ou de bloguer un article et bien sûr, forte contribution de l'internaute. Sur certains sites, l'étiquetage des articles via le système de tags est proposé.

L'intérêt est de pouvoir positionner certaines informations les unes par rapport aux autres et connaître la popularité d'une actualité ou d'un sujet (entreprise, personne) abordé (*la critique de mon produit est-elle très "populaire" ?*). C'est une approche alternative aux agrégateurs de news de type *Google Actualités* ou *Yahoo! Actualités* qui hiérarchisent les news avec de puissants algorithmes, mais sans recourir à l'humain. Ainsi, *Google Actualités* précise : "*La sélection et le positionnement des articles de cette page ont été réalisés automatiquement par un programme informatique*".

En outre, sur ces services de type *Digg*, l'information n'est pas brute : l'internaute qui poste l'actualité, peut, selon son expertise, y apporter un commentaire de valeur. L'information peut être commentée (en plus du vote) par les autres internautes. Comme de nombreux blogueurs, les utilisateurs des *Digg Like* jouent aux journalistes ou tout simplement donnent leur opinion.

Lorsqu'il s'agit d'actualités concernant le produit ou le service d'une société, ces commentaires, positifs ou non, sont une véritable manne pour de nombreux professionnels de l'information : veilleurs, marketeurs, chargés d'études et autres analystes qui vont pouvoir recueillir, à moindre frais, des informations sur leur nouveau produit, leurs clients ou ceux de la concurrence.

Digg est l'un des sites les plus copiés du Web 2.0. S'il reste la référence, d'autres applications, très largement inspirées, tirent leur épingle du jeu. En France, *Scoopeo*, *Fuzz* et *TapeMoi* offrent eux aussi des services d'actualités collaboratives. On peut également citer *Rojo*, le canadien *Niouz*, l'espagnol *Meneames*, le réseau de sites *BlogMemes*⁵⁶. Même *Netscape*, pionnier du navigateur web, a lancé un portail s'inspirant du système de notation de *Digg*.

On peut reprocher à ce type de site de favoriser les catégories "techniques" et scientifiques, faisant la part belle aux contributions des *geeks* et autres accros du net, grands utilisateurs de services Web 2. Toutefois, les sites comme *Blogmemes* ou *Netscape* diffusent davantage d'actualités à teneur économique et sociale.

- Les mix actualités / contributions des internautes

*Wikio*⁵⁷ est un service d'origine française (créé par le fondateur de *Kelkoo*), à l'intersection de *Google Actualités* et de *Digg*. En effet, *Wikio* mixe les articles issus de sites d'actualités, de blogs avec des contributions des internautes. On peut prendre connaissance des informations de différentes manières :

- En naviguant : via les thématiques (France, International, Politique, Economie Finance, Santé, Science ...) ou via le nuage de tags dynamiques qui se modifient en fonction de la rubrique ou vous vous situez.
- En recherchant via le moteur
- En s'abonnant aux flux RSS : à chaque catégorie et à chaque tag correspond son flux RSS. De même, lorsque vous lancez une requête via le moteur, vous pourrez surveiller ces résultats via un flux RSS.

- les plateformes de contributeurs indépendants.

La défiance et la volonté d'indépendance vis-à-vis des grands médias traditionnels de la presse écrite et audio-visuelle se sont traduites dès 2001 par la création massive de blogs. Les blogueurs sont-ils et peuvent-ils s'improviser journalistes ? Ce débat, qui promet d'être sans fin est intéressant. De nombreux blogueurs (mais pas la majorité) sont devenus de véritables experts dans leur domaine, à force d'écriture, de lecture et d'analyses quotidiennes. Ainsi, les "mainstream medias" invitent maintenant dans leurs colonnes et plateaux les blogueurs les plus talentueux. Ces blogueurs experts ou professionnels ne constituent pas les seuls créateurs d'actualités "alternatives".

Des initiatives de plateformes d'actualités écrites par des "non professionnels" ont aussi vu le jour : *Agoravox*⁵⁸ est l'un des premiers portails de "journalisme citoyen", lancé par Joël de Rosnay et Carlo Revelli⁵⁹ en mai 2005. *Agoravox* permet aux internautes de proposer leurs articles, pour commenter, réfléchir à des sujets d'actualités si possible avec une approche et un regard différents des médias traditionnels. Le format des contributions est multimédia. Le site explique : "*Tout le monde peut devenir rédacteur d'AgoraVox. Il n'est pas nécessaire de savoir écrire avec un style journalistique affirmé. Chacun d'entre nous peut devenir une sorte de "capteur en temps réel" de ce qui se passe dans notre entourage en fournissant des articles, des images, des extraits vidéo ou audio.*" "*Ce qui est important c'est de pouvoir témoigner d'un événement auquel vous avez assisté ou que vous avez découvert en faisant des recherches*". Il poursuit : "*Pourquoi devenir rédacteur ? Tout simplement pour faire entendre votre voix et pour relater des faits que vous avez identifiés et qui ne vous paraissent pas être suffisamment repris dans les médias traditionnels voire ignorés ou censurés.*"

Agoravox encourage donc à utiliser son style personnel. Il convient de s'inscrire pour devenir rédacteur. Les articles doivent toutefois obéir à une politique éditoriale⁶⁰. Un comité de rédaction valide ensuite les contributions.

En quelques mois, *Agoravox* a prouvé que de "simples" internautes pouvaient être aussi bon rédacteurs et "experts", en fonction de leurs profession, passion et centres d'intérêt, à l'instar des blogueurs professionnels.

Agoravox n'a pas la vocation d'un agrégateur d'actualités de type *Digg*, *Wikio* ou *Google News*. C'est une sorte de quotidien qui privilégie les articles de fond et l'analyse. Bien sûr, des rédacteurs "stars" écrivent régulièrement, et leurs articles sont attendus. Mais ce ne sont pas les rubricheurs traditionnels des autres médias.

Comme sur les autres sites Web 2.0, le lecteur peut attribuer une note d'intérêt à chaque article et poster un commentaire. Enfin, l'étonnante technologie *ReadSpeaker* ⁶¹ permet de retranscrire automatiquement tous les textes en version audio !

Agoravox n'est pas le seul exemple de site de journalisme "citoyen" : Le portail *OhMyNews* ⁶², initiative coréenne, est basé sur le même principe. Lancé en février 2000, le réseau international *OhMyNews* réunit plus de 40 000 rédacteurs.

Pour le veilleur, ces sites d'actualités collaboratives constituent des sources d'informations alternatives, avec une partie de l'actualité issue des organes et agences de presse et l'autre partie provenant des internautes. D'autre part, ces nouveaux médias permettent de découvrir et surveiller des "experts" voire des leaders d'opinion.

The screenshot shows the Agoravox website interface. At the top left is the logo 'AGORA VOX Le média citoyen'. To the right is a login section for 'ESPACE RÉDACTEUR' with a 'Login' field and an 'OK' button. Below the login is a 'login oublié ?' link. Further right is a 'Participez à la rédaction d'Agoravox' section with a 'DEVENIR RÉDACTEUR' button and a 'S'INSCRIRE' link. Below this is a navigation menu with categories: ACTUALITE, DOSSIERS, EDITION DU JOUR, QUI SOMMES NOUS?, PUBLIER, and RECHERCHE. A secondary menu lists topics: CITOYENNETÉ - ECONOMIE - ENVIRONNEMENT - EUROPE - INTERNATIONAL - MÉDIAS - POLITIQUE - SANTÉ - SOCIÉTÉ - TRIBUNE LIBRE. The main content area is titled 'ENVIRONNEMENT' and features an article by Jean Zin from Monday, May 14, 2007. The article title is 'La disparition des abeilles causée par les téléphones portables ?'. It includes a progress bar for 'Article intéressant ?' at 7% and a '93%' rating. There are 97 reactions and 75 votes. The article text discusses the disappearance of bees due to electromagnetic pollution from mobile phones. To the right of the article is a section for 'L'AUTEUR DE L'ARTICLE' featuring a profile picture of Jean Zin (Figeac), a member of the GRIT (transdisciplinary research), philosopher, and ecologist. At the bottom of the article, there is a small image of bees on a yellow flower.

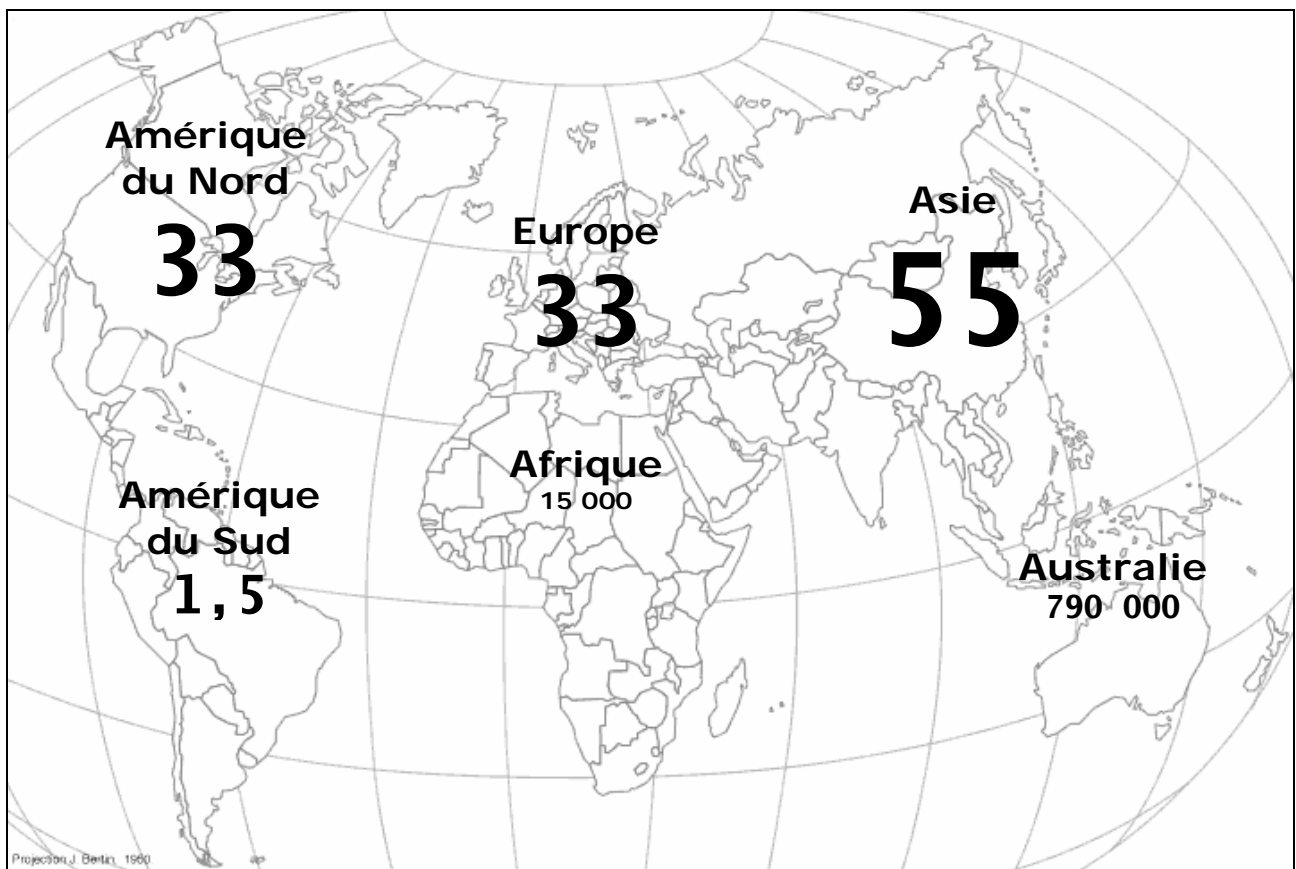
▲ Un article sur *Agoravox*. A droite, une présentation de son auteur. En haut, la possibilité de voter et commenter l'article.

II. Rechercher, Collecter

Dorénavant, tout internaute peut créer et publier sa propre information. L'identification des sources d'information et de l'information elle-même est devenue complexe car elles sont de plus en plus morcelées et atomisées. Il convient d'utiliser une large palette d'outils de recherche pour collecter les données en provenance des blogs, des bookmarks et réseaux sociaux.... Pour faciliter la collecte, il est en outre possible de verticaliser la recherche et de faire appel à des "humains". Par ailleurs, d'autres outils permettent d'archiver ses démarches de requêtes et de navigation.

a. La recherche de blogs et de flux RSS

Le volume de blogs est en constante croissance sur le net. Le nombre de blogs a littéralement explosé sur toute la planète. Estimé à 70 millions à l'été 2005, ils sont évalués à plus de 120 millions en mars 2007. *Technorati*, l'un des moteurs de recherche de blogs les plus utilisés mesure en 2007 un doublement du nombre de blogs tous les 320 jours. 120 000 blogs seraient ainsi créés chaque jour.⁶³



▲ Nombre de blogs en milliers et millions (la Russie est intégrée dans l'Asie). Estimation DIGIMIND Services ⁶⁴.

▪ **Les fils RSS omniprésents**

A chaque blog correspond son flux RSS, ce format qui permet d'extraire et de surveiller tout nouveau contenu qui paraît sur le blog. Si un vrai blog implique toujours un flux RSS, un flux RSS n'induit pas forcément un blog. En effet, le RSS et son concurrent Atom sont présents sur de plus en plus de type de médias et outils internet, dès lors que leur contenu est appelé à se modifier régulièrement.

On pourra donc trouver des flux RSS sur :

-les blogs

-les sites de presse et d'actualité. Exemple : *Les Echos, Libération, La Tribune, Le Figaro*)

-les sites de revues et magazines spécialisés. Exemple : La revue *Nature* propose ses fils RSS concernant ses publications spécialisées par domaines. *Science* fait de même⁶⁵.

-les sites corporate des entreprises : page d'actualités, catalogues produits. (Ex: Air Liquide)

-les résultats de moteurs de recherche de blogs. Tous proposent des flux RSS permettant de surveiller une requête (*Technorati, Google Blog Search, Feedster, IceRocket...*).

-les pages de résultats de moteur de recherche web. Certains de ces moteurs génèrent un flux RSS permettant de surveiller l'apparition de nouveaux résultats pour une requête donnée.

MSN Search (devenu *Live Search*⁶⁶) le proposait depuis janvier 2005, et chaque page de résultats offrait une icône orange en bas de page, permettant d'obtenir des fils RSS. Cette icône RSS a disparu sur *Live Search* mais les flux RSS sont toujours là (ils sont détectables automatiquement avec un navigateur *Firefox, Maxthon, IE7...* ou votre lecteur RSS *Attensa* ou *RSS Bandit*, par exemple).

Depuis peu, *Yahoo! Search Web* ⁶⁷ offre lui aussi un fils RSS sur ses résultats, là encore détectable via un lecteur ou navigateur (cela fonctionne uniquement pour le *Yahoo! US*).

-les moteurs d'actualités. La majorité des portails qui agrègent des flux d'actualités proposent des flux RSS. Exemple : *Google Actualités*, *Yahoo! Actualités*, *Topix*, *MSN Actualités*, *Wikio...* ⁶⁸

-les *Newsgroups Usenet* (groupes de discussion). Les fils RSS sont en effet aussi présents sur *Google Groupes* (ou *Google Groups*), ce qui rend plus aisé la surveillance des discussions. Chaque groupe (Google ou Usenet) affiche un fils RSS ou Atom ⁶⁹ (et ce quelque soit la langue) signalé par le logo orange RSS en bas de page;

-certains forums web. Exemple : *Forum Mobiles* ⁷⁰

-certaines listes de diffusion : elles peuvent être surveillées via l'abonnement à leur flux RSS. Cela évite l'engorgement de votre boîte mail. Exemple : la liste *Veille sur Yahoo! Groupes* ⁷¹ dispose d'un flux RSS.

-les bookmarks sociaux : *Del.icio.us*, *Yahoo! MyWeb2*, *Blogmarks* ⁷²....Le fils RSS permet d'être alerté de l'ajout de nouveaux favoris pour un tag ou un utilisateur définis.

-les plateformes de partage de photos. Exemple : *Flickr*, *PhotoBucket* ⁷³

-les services de partage de vidéos comme *YouTube*, *DailyMotion* ⁷⁴ pour être alertés des nouveautés.

-de très nombreux sites offrant un contenu qui se renouvelle.

The screenshot shows the LaTribune.fr website interface. At the top, there's a navigation bar with links like 'Ma page RSS', 'Ma bibliothèque', 'Mes alertes', etc. Below that, a search bar and a main menu with categories like 'Entreprises & secteurs', 'Bourse', 'Economie', etc. The main content area is titled 'Bourse / Actualité des sociétés en RSS' and features a list of companies with an 'RSS' link next to each name. A left sidebar contains various financial tools and information links.

Derniers articles au format RSS	
DALET	RSS
DAMARIS	RSS
DAMARTEX	RSS
DANE-ELEC MEMORY	RSS
DANIEL HARLANT	RSS
DANONE	RSS
DASSAULT AVIATION	RSS
DASSAULT SYSTEMES	RSS
DEBUSCHERE	RSS
DELACHAUX	RSS
DELFINING INDUSTRY	RSS
DELL	RSS
DELTA PLUS GROUP	RSS
DESQUENNE ET GIRAL	RSS
DEVEAUX	RSS
DEVERNOIS	RSS
DEVOTEAM	RSS
DEXIA	RSS
DIAGEO	RSS
DIDOT-BOTTIN	RSS
DIGIGRAM	RSS
DIGITECH	RSS
DIREKT ANLAGE BANK	RSS
DISTRIBORG GROUPE	RSS
DLSI	RSS
DMC	RSS
DIAGNOSTIC MEDICAL SYSTEMS	RSS
DOCKS LYONNAIS	RSS
DREAMNEX	RSS
DU PAREIL AU MEME	RSS
DUC	RSS
DUC LAMOTHE PARTICIPATIONS	RSS
DUPONT DE NEMOURS	RSS
DURAN	RSS
DYNACTION	RSS
DYNAFOND	RSS

▲ Veille Corporate : La Tribune propose des flux RSS segmentés par sociétés.

Le veilleur, via ces flux RSS, pourra surveiller de nombreuses sources d'informations. Par ailleurs, au-delà des blogs "introspectifs" d'adolescents dont peu présentent un intérêt pour les professionnels, certains types de blogs peuvent se révéler utiles :

- Le blog d'actualité et d'opinion présentant réflexions et critiques de journalistes et internautes,
- Le blog thématique édité par des passionnés, experts, professionnels, consultants et techniciens,
- Le blog d'entreprise, outil de communication externe ou interne à l'initiative des salariés ou des dirigeants.

▪ Les outils pour trouver des blogs et flux RSS

Sur ces moteurs, la recherche s'effectue par mots clés classiques ou par tags comme sur *Technorati* par exemple. Le contenu indexé (blogs et flux) est très hétérogène. Aussi, l'utilisation de la recherche avancée sur ces outils est souvent nécessaire. En effet, un blog peut avoir abordé la thématique "développement durable" | seule fois parmi ses 4500 billets. Il conviendra donc alors de centrer sa recherche sur le titre du blog ou son url.

Les moteurs :

Certains moteurs sont dédiés à la recherche de blogs. C'est en fait un abus de langage. Ces moteurs se focalisent sur la recherche de flux RSS mais ne peuvent faire la différence entre un flux provenant d'un grand journal et celui d'un blog. Ainsi, le moteur *Technorati* classe parmi les blogs les plus populaires...*Le Figaro*.

Pour mener une recherche efficace, il conviendra d'utiliser plusieurs moteurs. Car comme pour la recherche web où il ne faut pas se limiter à Google au risque de manquer près de 85% des résultats⁷⁵, la recherche sur les blogs nécessite l'emploi de plusieurs outils⁷⁶. En voici quelques-uns :

> *Google Blog Search/ Recherche de blogs* : Google applique sa technologie puissante aux flux RSS et blogs. L'interface est notamment disponible en français et en anglais. Le moteur est doté de fonctions de recherche avancée (recherche sur le titre, le billet, recherche par ancienneté...)

> *Ask Blog & Flux* : Ce moteur permet de rechercher parmi les blogs, les flux XML (RSS ou Atom). Et aussi de s'abonner à un blog ou un flux RSS.

> *Feedster* : Ce moteur recense plus de 900 000 sources et 5000 sont rajoutées chaque jour.

> *BlogPulse* : Indexe plus de 46 millions de blogs.

> *Technorati* : Ce moteur rafraîchit son index toutes les 5 à 15 minutes et recherche parmi plus de 70 millions de blogs (en avril 2007 ⁷⁷) ou sites proposant des flux RSS. Il est possible de rechercher par mots clés ou par tags.

> *Sphere*: va prendre en compte plusieurs variables pour mieux chercher parmi la blogosphère :

- comme pour le contenu web classique, les réseaux de liens entrants et sortants
- la qualité de ces liens

- les internautes à l'origine de ces liens (simple commentateurs ou leaders d'opinion)

3 moteurs web intègrent la recherche sur les blogs au sein de leur recherche avancée ⁷⁸ :

> *Yahoo! Search*

Pour centrer sa recherche sur les flux RSS et les blogs, il convient d'aller dans la recherche avancée (à droite de la boîte de requête) puis dans le menu déroulant "Format de fichiers", et de sélectionner le format RSS/XML.

> *Exalead*

Après avoir effectué votre recherche, il suffit de cliquer sur le bouton RSS jaune dans la fenêtre "Préciser la recherche". Au sein des résultats de recherche, le flux RSS est disponible via l'icône RSS. Simple et propre.

> *Microsoft / MSN / Live Search*

MSN a amélioré les interfaces de son moteur pour restreindre la recherche sur les fils RSS. Après l'affichage de votre page de résultats, vous pouvez affiner votre recherche grâce au bouton "Flux (ou "Feed" en anglais). Ce bouton est accessible via le menu "Plus". Au sein des résultats de recherche sont alors disponibles l'url du flux, le cache et un lien pour s'inscrire au fils RSS.

Au-delà de ces moteurs spécialisés, il peut être bénéfique de tenter une recherche sur un moteur Web (*Google* ou *Yahoo! Search* par exemple) de type : *blog OR RSS [votre sujet]*. Exemple : "*blogs sur les télécoms*". Cela vous permettra, non pas de chercher au sein des blogs mais de trouver d'autres blogs.

Les répertoires ⁷⁹

Comme pour le web, un certain nombre d'annuaires entendent recenser les blogs et fils RSS. Attention, comme la blogosphère est très mouvante et en constante croissance, ces guides sont par nature non exhaustifs et leur mise à jour régulière est difficile.

> *Blogarama* : Ce répertoire recense plus de 49 000 blogs classés par catégories

> *RSS Network* : indexe plus de 150 000 fils RSS

> *RSSFeeds* : Répertoire plus de 119 000 fils organisés en catégories.



b. La recherche par tags

Nous l'avons vu, avec la *Folksonomie*, l'internaute peut désormais mettre en place et partager sa propre classification via les tags. L'utilisateur affecte l'étiquette de son choix à ses favoris, images, vidéos. L'avantage est indéniable : c'est l'internaute qui définit son étiquette et non plus un organisme. Mais les inconvénients sont nombreux : difficile gestion des singuliers/pluriels, problème de synonymes, de signification (que désigne cet internaute par ce tag), d'homonymes, de polysémie, d'hétérogénéité : plusieurs tags peuvent désigner les mêmes documents.

Toutefois, l'approche par tag facilite la recherche, d'une part en ajoutant une couche sémantique aux documents web et d'autre part, en permettant une navigation par concepts, de plus en plus abandonnée compte tenu de la désertion progressive des catégories des grands répertoires web par les internautes.

De plus, on peut tagger pratiquement tout type de contenu web. Ainsi, les sites *43 Things* et son équivalent français *36 Trucs*⁸⁰ proposent aux internautes de lister leurs objectifs, expériences ou projets. Chacun de ces projets est taggé via un ou plusieurs mots clés

Les moteurs et métamoteurs de tags

Les moteurs

La plupart des applications de "social bookmarking" ou de partage de fichiers images ou vidéos permettent donc de tagger leur contenu. Sur certaines, le moteur de recherche interne rend possible, en plus de la recherche sur le *full text*, une recherche centrée sur les tags. C'est le cas sur *Flickr* : vous pouvez rechercher, au choix, sur le titre et la description des images ou sur les tags associés, définis par les utilisateurs qui ont placé leurs photos.

Un outil comme *Del.icio.us tag search*⁸¹ va vous permettre d'effectuer des recherches via un ou plusieurs tags sur le service de social bookmarking *Del.icio.us*.

Pour la recherche de blogs ou flux RSS, le moteur *Technorati* propose une recherche classique sur les mots clés contenus dans les billets mais aussi une

recherche par tags⁸². En effet, l'éditeur d'un blog peut tagger chacun de ses billets via des mots clés qui les catégorisent. Ces mots ou tags seront repris par l'index du moteur. Ainsi, la recherche du mot *pollution* parmi les tags affiche un billet sur le réchauffement climatique. Si ce billet est bien catégorisé par son auteur dans une rubrique "pollution" (parmi d'autres), il ne contient pas ce mot dans son texte, toutefois en relation avec le concept de pollution. Clairement, une recherche classique sur le seul texte des billets n'aurait pas ramené ce billet ou du moins, pas dans les premiers résultats. Autre intérêt : *Technorati* vous propose des tags associés. Ainsi pour *pollution*, il affiche *environment, politics, global warming, science, china, energy, news, air, climate, climate change*, autant de mots clés pour élargir ou affiner vos requêtes.

Le moteur *SearchLess*⁸³ offre une approche originale : il associe la recherche par mots clés à la recherche par tags. Ainsi, si vous effectuez une requête avec le mot "intelligence", vous pouvez préciser par le tag "CIA". Les résultats sélectionneront donc le mot *intelligence* sur des contenus (billets de blogs par exemple) taggés *CIA* par l'utilisateur lui-même : on est alors sûr que les 2 concepts sont intimement liés.

Métamoteurs

Des métamoteurs⁸⁴ permettent de rechercher sur les différents types de contenu "sociaux" : bookmarks, images, vidéos.

Ainsi, *Tag Central* agrège de nombreux formats web "taggés". En lui indiquant un tag, il affiche des résultats provenant à la fois de *Flickr, Smugmug* (photos), *Technorati* (blogs), *Blogmarks, Connotea, Furl, del.icio.us* (bookmarks sociaux), ou encore *43Things* (listes de projets).

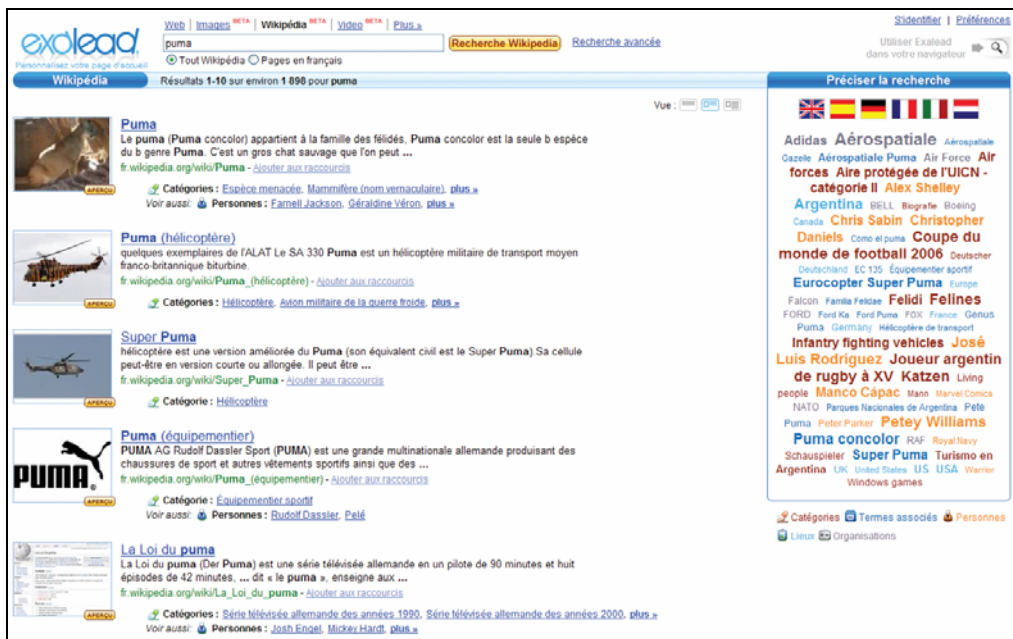
Zewol, outil de "méta syndication francophone" permet de rechercher parmi les tags (*Flickr, del.icio.us...*) mais aussi sur les actualités (*Google News, Ask Actualités, Yahoo! Actualités...*), les blogs (*Technorati, IceRocket, Google Blog Search, Ask Blogs...*), le "buzz" (*Tape.moi, Scoopeo, Wikipedia...*), les vidéos (*Google Videos, Podscope Video, Daily Motion...*) et un "mix" (*Furl, Echonimo actualités, Plazoo, Feed24*). Concrètement, c'est une bonne approche de recherche globale sur de nombreuses ressources de type Web 2.0

Il est toujours possible de traquer la présence ou l'apparition d'un mot clef sur le web via la mise en surveillance du flux RSS de la requête.

Ketoag est un outil qui permet d'interroger de nombreux services sociaux

(Del.icio.us, Furl, Diig, Wikio, Blinklist) mais également des moteurs de recherche (Google Blog Search). Par ailleurs, vous pouvez générer, pour chaque requête, un flux RSS mais aussi un fichier OPML (format permettant l'agrégation et l'échange de listes de flux RSS).

Gnosh est un métamoteur qui recherche parmi les moteurs web, les moteurs de blogs et les outils de partage de favoris. De plus, une recherche sur l'onglet "People" vous permet de rechercher parmi les favoris des internautes inscrits à la communauté Gnosh.



▲ Exalead recherche sur Wikipedia. A droite, le nuage de tags.

Les nuages de tags

Ces "nuages" regroupent un ensemble de tags dans un même espace en affectant à chacun une taille variant selon leur popularité (fréquence d'utilisation ou de requête sur ces mots) ou la densité d'articles auxquels ils se réfèrent (dépêches d'actualités par exemple).

Les nuages de tags sont devenus à la mode car ils permettent une navigation facile, très visuelle, au sein d'un site, d'une base de donnée ou des résultats d'un moteur. Ils apportent une grille de lecture supplémentaire à des sites dont les menus sont souvent surchargés ou le contenu atomisé.

Nombreuses sont les applications qui utilisent le nuage de tags pour optimiser la navigation et mettre en valeur leur contenu : le partage de vidéos (DailyMotion), les bookmarks sociaux (*Del.icio.us*, *MyWeb2* de Yahoo!) ou les services d'actualités collaboratives comme Wikio. Certains moteurs comme *Exalead* ont également choisis d'afficher des nuages de tags au sein de leurs pages de résultats. Dans ce cas, il s'agit d'une aide à l'affinage des requêtes : un clic sur un tag va afficher une nouvelle page de résultat prenant ce mot en compte.

*Zibb*⁸⁵, moteur de recherche vertical dédié au Business to Business utilise aussi le nuage de tags pour visualiser clairement les mots clés de thématiques connexes les plus fréquents à votre requête.

D'autres sites plus classiques ont également choisi l'approche du nuage de tag pour rendre leur navigation plus attractive et leurs rubriques plus visibles. Ainsi, le site *Ecrans*⁸⁶ (du journal *Libération*) affiche en bas à droite un nuage de tags mettant en valeur son contenu très touffu.

Certains outils proposent de générer des nuages de tags pour des sites ou moteurs n'en disposant pas :

*Google Tag Cloud Maker*⁸⁷ est un script pour *Firefox* qui permet d'afficher un nuage de tags à droite des résultats de Google : plus un mot est fréquent dans les résultats de votre recherche, plus la taille d'un tag est importante.

*Newzingo*⁸⁸ extrait des mots de *Google News* pour constituer un nuage d'occurrences parmi les termes de l'actualité. Plus le terme est gros, plus les articles relatifs à ce terme sont nombreux. *Techzingo* réalise le même process mais cette fois, les cibles sont des médias de référence dans le High Tech et le logiciel : *Slashdot*, *Digg.com* et *TheRegister*.

The screenshot shows the IMDb website interface. At the top, there are navigation tabs for 'NOW PLAYING', 'MOVIE/TV NEWS', 'MY MOVIES', 'NEW ON DVD', 'IMDb TV', 'MESSAGE BOARDS', 'SHOWTIMES & TICKETS', 'IMDb pro', and 'IMDb Resume'. Below these is a search bar with the text 'search: All' and a 'go' button. To the right of the search bar is a 'Keywords:' section with checkboxes for 'Based On Novel' and 'Independent Film'. Below this is a 'Related Keywords' section with a list of terms like '1900s', '1910s', 'Accident', 'Adultery', etc. To the right of the related keywords is a 'MOKA Movie Keyword Analyzer' section with a search box and a 'Go' button. Below the MOKA section is a list of movie titles with their release years and ratings, such as 'Maynila, Sa mga kuko ng lwanag (1975)' with a rating of 9.0/10.

▲ Un nuage de tags sur l'Internet Movie Database : une manière de naviguer parmi tous les mots clés associés aux films.

La qualification du contenu par les utilisateurs ("folksonomie") est une réelle valeur ajoutée. Les moteurs de recherche font reposer une grande part de l'effort sur l'utilisateur final : une page de résultats de Google doit être évaluée, souvent interprétée par l'internaute. La folksonomie fait, elle, reposer l'effort de filtrage sur l'utilisateur initial, celui qui va rechercher et appliquer ses propres "tags". Parce qu'en amont 1743 personnes ont taggé une page avec un groupe de mots clés, l'utilisateur final, lors de sa requête, reçoit une page de résultats souvent très pertinente.

Sur un moteur classique, c'est à l'utilisateur d'imaginer ses mots clés pour affiner sa recherche. Sur les bookmarks sociaux de type *Del.icio.us*, les internautes qui partagent leurs ressources ont déjà préparé ce travail d'affinage via leurs tags. Prenons le cas d'une recherche "extrême" car générant de la synonymie : la requête "Jaguar". Sur Google.com, on me propose certes des mots clés connexes pour affiner mais tous sont liés à la voiture. Sur *Del.icio.us*, les tags affichés sont relatifs à la voiture mais aussi au système d'exploitation, à un groupe de musique et à l'animal.

Pouvoir rechercher parmi les tags, classifications émanant de l'utilisateur, peut être particulièrement intéressant lorsqu'une recherche classique par mots clés sur le contenu des documents s'est avérée infructueuse ou génère des résultats peu pertinents.

Le tagging, une pratique quotidienne ?

Selon une enquête menée en décembre 2006 par le Pew Internet & American Life Project, 28% d'utilisateurs des internautes américains ont taggé ou catégorisé en ligne des contenus tels que des photos, des billets dans des blogs, des articles d'actualité... Sur une base journalière, 7 % des utilisateurs d'Internet pratiquent le "tag" en ligne.

Ce système de « taxonomie populaire » pourrait-il constituer une alternative aux outils de recherche traditionnels par mots clés ? Il convient de distinguer les usages professionnels des usages grand public. Olivier Le Deuff⁸⁹ reprend, dans son étude sur les "Folksonomies", la distinction effectuée par Thomas Vander Wal entre les folksonomies "étroites" (narrow folksonomies) et "générales" (broad folksonomies). Selon Olivier Le Deuff, "cette distinction est importante, car les premières sont surtout utilisées dans un objectif individuel tandis que les secondes privilégient l'aspect collectif et collaboratif du partage d'information. Ainsi les sites de partages de favoris, comme *del.icio.us* ou *Connotea*, sont plutôt des folksonomies générales puisqu'un même site peut être partagé par plusieurs utilisateurs et recevoir le même tag. Ce type de folksonomies s'appuyant sur des

réseaux sociaux ne fait pas que classer de l'information et la partager. Il met en relation des usagers qui partagent les mêmes centres d'intérêt. L'utilisateur indexeur devient à son tour un peu indexé et mis en relation à la fois avec d'autres mots-clés, d'autres sites et d'autres usagers".

Il est clair que dans le domaine de l'information professionnelle ou spécialisée, le manque d'uniformisation des folksonomies, qu'elles soient générales ou à vocation plus étroite, pose les limites de ce système. Les outils d'indexation et de classification traditionnels ont en effet été créés pour des corpus structurés, aux limites clairement définies, et s'adressent à des utilisateurs experts d'un domaine ou de la recherche d'information. L'absence de hiérarchisation et de contrôle constitue à la fois le point fort et le point faible de la folksonomie. La barrière des langues pourra constituer également un obstacle, les tags étant parfois mélangés dans un même site sans distinction. De plus, comme le note Cyril Fievet ⁹⁰, la pratique du tagging n'est pas uniforme d'un site à l'autre. "Sur *Technorati*, les tags indexés sont ceux choisis par les blogueurs et fonction du contenu qu'ils ont eux-mêmes publié. Sur *del.icio.us*, en revanche, ce sont les internautes qui taguent le contenu des pages qu'ils conservent dans leurs signets ».

Il convient également d'être vigilant quant à une utilisation « frauduleuse » des tags, destinée uniquement à garantir la popularité d'un blog comme cela a pu être le cas avec les méta-données incorporées dans le code source d'une page html.

Élie Francis et Odile Quesnel présentent dans la revue "Documentaliste, Sciences de l'information" les quatre modes d'indexation et de classification sur le web : la classification personnelle, l'indexation par l'auteur, l'indexation par l'utilisateur et la classification globale. Et s'interrogent sur l'avenir de l'indexation collaborative, à la croisée de la qualité (liberté, diversité, enrichissement) et de la désinformation (subjectivité, manque de cohérence).

Il n'en reste pas moins que folksonomie et indexation classique ne s'opposent pas nécessairement, et constituent des moyens de description de contenus adaptés à des contextes différents. Stephen Lee, créateur du site *Guten Tag* voit dans la folksonomie une sorte de "web sémantique du pauvre", peu onéreux à mettre en place, mais dont le succès reposera sur une série de bonnes pratiques collectives. L'intérêt de ce système résidera non pas dans la capacité à trouver une réponse à une question précise, mais plutôt à découvrir des sources originales et peu connues, ou de localiser des experts d'un domaine pointu. Les tags les plus rares et précis pourront entre autres trouver leur application dans le domaine de la veille, mettant

en œuvre la "sérendipité", ou l'art de favoriser le hasard pour des découvertes heureuses.

Laissons la conclusion sur ce point à l'Association des Professionnels de l'Information et de la Documentation (ADBS), qui estime dans sa newsletter que "Le foisonnement des propositions de mots-clés aboutirait à un résultat plus efficace qu'une indexation réalisée par un vocabulaire contraint. Cette affirmation semble être démontrée... pour l'instant".

c. La recherche de personnes et d'informations entreprises

La recherche de données sur les personnes physiques ou morales s'effectuait jusqu'ici principalement en interrogeant les bases de données d'informations légales et financières, gratuites ou payantes. L'interrogation de service tel que *Societe.com*, *Infogreffe* ou *Yahoo! Finance* est bien sûr encore nécessaire. Mais il faut compter maintenant avec les réseaux sociaux professionnels qui eux aussi, rassemblent des données sur les entreprises ou les individus.

L'intérêt des réseaux sociaux

Que peuvent-ils apporter par rapport aux services classiques d'informations d'entreprises ? D'abord, comme expliqué précédemment, ils sont ouverts à tous. Les informations y figurant ne sont plus le seul fait de l'éditeur du site et ne sont pas uniquement constituées par les fichiers de données légales publiées par les organismes officiels comme le Greffe du Tribunal de Commerce en France. Sur les réseaux sociaux centrés sur les personnes, le contenu est intégralement amené par l'internaute qui affiche son CV détaillé, ses publications, ses œuvres (textes ou multimédias). L'intérêt est évident : des personnes peu visibles auparavant parce que ne faisant pas partie des dirigeants d'entreprises (gérant, fondateur, membre du conseil) sont maintenant accessibles et donc "requêtables" par les internautes. On peut donc collecter des informations sur un cadre de R&D d'une société de télécoms concurrente ou sur le management étonnant de telle entreprise de plasturgie, citée dans le CV d'un employé "de passage"...

A l'instar des blogs, l'information présente sur les réseaux sociaux ne tient plus du discours ou des données officielles puisque c'est l'internaute qui communique sur sa vie et son entreprise.

Gérer l'éparpillement de l'information

D'autres plateformes de social networking mélangent l'information collectée sur le web avec les contributions des utilisateurs (à l'instar de *Wikio* pour les actualités). Ainsi, le service *ZoomInfo*⁹¹ permet de rechercher des sociétés et des personnes en parcourant des milliers de données extraites du web mais aussi en provenance de profils d'individus et d'entreprises renseignés par les sociétés ou les employés eux-mêmes. Aussi, cohabitent des informations sur des entreprises cotées avec des données de firmes non cotées, des fiches de cadres du top management et des conseils d'administration avec des CV d'employés plus "fonctionnels".

Le français *Ziki*⁹² utilise aussi des informations collectées sur le web : c'est une sorte d'annuaire perfectionné qui indexe et archive non pas les données signalétiques mais les contenus produits par une personne à partir du moment où ils disposent d'un flux RSS (billets de blogs, photo Flickr, favoris Del.icio.us,...). *Ziki* ne permet pas de créer du contenu propre mais agrège des données élaborées par un même individu, auparavant éparpillées sur l'internet. Bref, vous améliorez votre visibilité et facilitez la recherche de vos productions sur le net en créant un point unique d'accès.

La recherche sur tous ces réseaux sociaux implique souvent l'interrogation fastidieuse, après inscription, de plusieurs dizaines de sites.

Aussi, certains outils⁹³ proposent une métarecherche en lançant une même requête sur plusieurs services de "social networking" : *Stalkerati* recherche le nom d'une personne sur les réseaux Friendster et Facebook. Plus complet et ciblant davantage les réseaux professionnels, *YoName* recherche les noms sur *MySpace*, *LinkedIn*, *Friendster*, *Windows Live Spaces*, *Bebo*, *Digg* ou *Xanga*.

Le moteur de recherche de personnes *Wink* scrute les réseaux sociaux *MySpace*, *Bebo*, *Friendster*, *LinkedIn*, *Live Spaces* et autres. De plus, avec cet outil, il est possible de créer son propre profil et de le connecter à des réseaux sociaux.

Attention toutefois, si vous êtes un utilisateur recherchant des données en français, la nature des résultats est plutôt tournée vers le monde anglophone.

Ces réseaux sociaux ne sont pas suffisants pour une recherche professionnelle complète. A l'instar des blogs face à la presse traditionnelle, ils constituent un complément intéressant aux bases de données officielles.

d. Les moteurs de recherche "humains"

Après être apparus sur le net à la fin des années 90, les services de questions-réponses avaient laissé la place aux algorithmes tout puissants des moteurs de recherche de type Google. Mais les robots et programmes ont tous besoin d'une intervention humaine pour gérer la qualité et la pertinence des résultats.

Ainsi, tous les grands moteurs majeurs ne font pas confiance à leurs seuls algorithmes. Les employés de Yahoo! sont mis à contribution pour surveiller la qualité des résultats fournis par le moteur. Chez Google et MSN Search, des "process" d'évaluation humaine avaient été identifiés à travers les sites *Eval.Google.com* et *MSN HRS* notamment.

La revanche de l'humain

Au-delà de ce "rating" humain sur les résultats de moteurs de recherche, de nombreux services, depuis 2 ans, mettent à disposition des "experts" en chair et en os. C'est alors un véritable mode collaboratif qui se met en place lorsque le produit est bien pensé. C'est la cas de *Yahoo! Answers* ("Yahoo! Questions Réponses" en France) ⁹⁴. Ce service gratuit de questions réponses lancé fin 2005, permet à l'internaute de poser une question dans un formulaire. Chaque membre de la communauté Yahoo! (à savoir tout ceux qui disposent d'un compte Yahoo!) peut jouer le rôle d'"expert". Yahoo! propose aussi le service (également gratuit) *Ask Yahoo!* qui propose de répondre à vos questions mais cette fois via un groupe d'"experts" sélectionnés par Yahoo! Inc. *Yahoo! Answers* a certes une orientation initiale grand public mais des questions un peu techniques voire scientifiques trouvent réponse. En effet, parmi tous les contributeurs à *Yahoo! Answers*, figurent des métiers et expertises très diverses, du fait de son audience importante. Il y cohabite des types de questions variés : "*Quels sont les "groupes stratégiques" de l'industrie mondiale du jouet ?*", "*Quelles sont les données économiques du marché de l'assurance vie collective en France ?*"...

Autre point intéressant : le service est déployé dans de nombreux pays (Etats-Unis, Royaume-Uni, Canada, Australie, Inde, Espagne, Brésil, Argentine, Mexique, Italie, Allemagne, Singapour, Taiwan, Corée du Sud, Hong Kong, Japon, Chine...). Aussi, il est possible de "localiser" certaines questions et de bénéficier parfois d'une expertise pays.

L'une des raisons du succès du service de Yahoo! tient au fait de baser les questions et réponses sur la contribution d'internautes volontaires. Chacun participe, constitue un vrai réseau et peut surtout se valoriser en devenant "expert" dans son

domaine de compétences ou sa passion. Il est même possible de créer sa propre communauté qui rassemblera vos amis mais aussi vos "experts" préférés qui répondent à vos questions et centres d'intérêt. Le pool de répondants n'est donc pas limité à une liste restreinte et il s'auto-entretient. D'autre part, les "experts" s'évaluent entre eux et sélectionnent parmi une liste, si nécessaire, ce qu'ils considèrent comme la meilleure réponse à une question.

The screenshot shows the Yahoo! Answers interface. At the top, there's a navigation bar with the Yahoo! logo, the word 'ANSWERS', and a user greeting 'Welcome, christophe' with links for 'Sign Out' and 'My Account'. Below this is a search bar with 'ask.' and 'answer.' buttons, and a green banner that says 'be a better PLANET find out HOW'. The main content area is titled 'Resolved Question' and features a question: 'Market research info on the urban/hip hop clothing market?' asked by user 'peace7'. The question text reads: 'I have a socially conscious based clothing line targeting the urban market. I am in the business planing phase and stuck on the market research part. the last report i read , this is a 400million a year market, but that dont make sense because Roca Wear alone did over 700million in sales in 2006... im trying to find something free. thanks for any help.' Below the question, there are social sharing icons and a 'Show who starred' link. A 'Best Answer' is provided by user 'Ivan', who states: 'This question is a little bit difficult to answer as you have not provided the country you are targeting. Or are you targeting the'.

▲ Une question sur Yahoo! Answers.

Sur ce marché des Questions-Réponses, *Google* a échoué, notamment parce qu'il a séparé formellement demandeurs et contributeurs, limitant ainsi fortement l'aspect collaboratif. Le Service *Google Answers*⁹⁵ permettait de poser des questions à des experts. Mais là où des services concurrents comme *Yahoo! Answers*, *Lycos IQ*,

*Windows Live QnA*⁹⁶ sont gratuits, *Google Answers* était un service payant, faisant appel à un réseau de 800 experts. Aussi, *Google Answers*, après 4 ans d'activité, a fermé. Les questions ne sont désormais plus acceptées mais l'accès aux archives des questions-réponses est maintenu.

Par ailleurs, les réseaux sociaux évoqués plus haut ont bien compris la puissance des communautés pour faire office de spécialistes. Ainsi, avec *LinkedIn Answers*, la plateforme de réseaux sociaux *LinkedIn* s'est lancée elle aussi dans le service de Questions/Réponses.

Ne pas oublier les services Web 1.0

Lorsque l'on parle de Web 2.0, on évoque rarement ces services de Questions-Réponses au profit de sites plus technologiques. Pourtant, nous sommes ici dans la logique des bookmarks sociaux : partager ses connaissances, faire bénéficier les autres de son savoir et en retour, profiter de l'expertise de la communauté. Comme la plupart des applications Web 2, la seule utilisation des services Q/R ne suffit pas. Elle doit venir en complément de recherches parallèles via d'autres outils. Ainsi, pour la veille, l'inscription à des listes de diffusion permet aussi de poser des questions à des experts ou passionnés. De plus, il subsiste des services d'expert "traditionnels", qui, s'ils ne mettent pas l'accent sur la collaboration, permettent d'identifier des compétences parfois très pointues. Certaines ressources répertorient des milliers d'experts qu'il faudra ensuite contacter personnellement pour bénéficier de leur conseil. C'est le cas du site *Experts.com* qui identifie des experts dans des domaines allant de l'acoustique à la biochimie en passant par l'industrie pétrolière⁹⁷.

e. les moteurs de recherche "collaboratifs"

Nous l'avons vu, certains services d'actualités collaboratives proposent des résultats issus d'un mixage entre contribution des internautes et collecte de nouvelles sur le web. Cette approche de mixité se met peu à peu en place pour les moteurs de recherche web. En effet, les modèles algorithmiques ou humains présentent tout 2 leurs limites en terme de pertinence, de fiabilité et d'exhaustivité.

Aussi, de plus en plus de projet de moteur mixtes, à savoir proposant des résultats web conjointement à des résultats et/ou sélections issus d'humains voient le jour. Du web donc, mais avec une forte dose de cerveaux humains.

Wikiasari

Celui qui fait le plus parler de lui est le projet de Jimmy Wales, le co-fondateur de l'encyclopédie Wikipedia. Baptisé *Wikiasari*⁹⁸ (combinaison de *wiki*, "rapide" en hawaïen et *asari*, "recherche fouillée" en japonais), le projet vise à créer un outil de recherche où les internautes pourront influencer sur les résultats en notant leur pertinence et/ou en en proposant de nouveaux, selon leur expertise.

Ce type de moteur s'inscrit donc à l'opposé d'outils comme Google aux algorithmes automatiques, dont les robots sont souvent trompés par des manipulations de webmasters (*spamdexing* par exemple).

Cet outil qui devrait voir le jour courant 2007 s'appuie sur des infrastructures Open Source existantes comme *Nutch* ou *Lucent*.

C'est un projet particulièrement intéressant : Comme sur l'encyclopédie ouverte Wikipedia où des milliers de contributeurs construisent et alimentent l'ouvrage, sur ce moteur, des réseaux de centaines d'internautes pourront apporter leur savoir pour améliorer la pertinence des résultats. Ce n'est plus seulement la puissance de calcul de machines qui sera mise à contribution mais aussi la puissance de la communauté.

Derrière le projet *Wikiasari*, il y a de gros moyens, financiers et humains, l'expérience de la fondation Wikipedia et le potentiel de trafic que pourra apporter l'encyclopédie ouverte : Fin 2006, Comscore a en effet classé Wikipedia comme le 6ème site le plus visité au monde avec 171,9 millions de visiteurs uniques au mois de novembre. On peut donc y croire.

Avec yoono, trouvez les sites, les fils d'informations et les experts sur vos sujets favoris

Accueil En savoir plus Recherche Application windows Extension firefox Blog Forum Aide

Rechercher : OK ?

Les sites similaires

- [Futura-sciences](http://www.futura-sciences.com/)
http://www.futura-sciences.com/ - (popularité: 32)
Suggestions sur cette page
- [Infoscience.fr](http://www.infoscience.fr/index.php3)
http://www.infoscience.fr/index.php3 - (popularité: 29)
Suggestions sur cette page
- [Larecherche.fr](http://www.larecherche.fr/)
http://www.larecherche.fr/ - (popularité: 25)
Suggestions sur cette page
- [Cybersciences](http://www.cybersciences.com/)
http://www.cybersciences.com/ - (popularité: 23)
Suggestions sur cette page
- [Techno-science.net](http://www.techno-science.net/)
http://www.techno-science.net/ - (popularité: 17)
Suggestions sur cette page
- [Cnrs.fr](http://www.cnrs.fr/)
http://www.cnrs.fr/ - (popularité: 15)
Suggestions sur cette page
- [Cybersciences](#)

Un expert, c'est quoi?
Un expert est un utilisateur qui publie certains de ses dossiers pour permettre à la communauté yoono de l'identifier.
Devenir expert permet d'entrer en relation avec les internautes qui partagent les mêmes centres d'intérêt.
Le nombre d'abonnés au dossier d'un expert permet d'évaluer son niveau d'expertise sur le sujet.

Nombre de résultats :
les 20 premiers

▲ Une recherche de sites similaires sur Yoono.

Decipho, Yoono

Certes, l'idée d'un moteur de recherche collaboratif n'est pas nouvelle. Récemment, on peut citer le lancement du moteur *Decipho*, ou en mars 2006 celui de Yoono. Même si les approches diffèrent, ces outils font tous la part belle à l'apport de chaque utilisateur dans le traitement des résultats.

*Decipho*⁹⁹ est un moteur basé sur la technologie de *Ask.com*, l'un des 4 moteurs les plus utilisés aux USA. L'internaute crée ses propres catégories de résultats partageables avec les autres. Il peut également tagger les résultats afin d'étiqueter un site. Aussi, lors d'une requête, on peut préciser les mots clés classiques avec des tags. On bénéficie alors de tous les tags mis en place par la communauté des utilisateurs. Un curseur permet de définir la proportion de catégorisation des autres utilisateurs à prendre en compte dans sa recherche.

Le français Yoono¹⁰⁰ propose une autre démarche encore : l'outil fonde ses résultats sur la mutualisation des connaissances des internautes par le partage des favoris. Le

moteur utilise les favoris que sa communauté d'utilisateurs a accepté de partager. Lors d'une requête, les résultats sont classés par popularité, à savoir selon le nombre de fois où une page a été mise en favoris par les utilisateurs. Une logique proche de *Del.icio.us* mais avec *Yoono*, l'ajout de favoris est automatique. De plus, l'outil affiche les utilisateurs, qui comme vous, ont référencé des sites sur la politique vénézuelienne par exemple. De cette manière, vous pouvez identifier des passionnés ou experts partageant vos domaines de prédilections. Les résultats ne sont plus basés sur les seuls algorithmes mais sur l'expérience et la capacitation des autres utilisateurs.

Les initiatives de moteurs collaboratifs, plus ou moins intéressantes sont multiples. *Foxx* est un moteur basé sur des communautés d'intérêt, et sur *Webzzle*, l'utilisateur précise les concepts qui l'intéressent afin de contextualiser les mots-clés choisis. ¹⁰¹

Humains ou algorithmes ?

Certes, le mode collaboratif ne remplacera pas totalement la recherche algorithmique. Les 2 sont complémentaires. L'une permet d'améliorer la pertinence, de prendre en compte des concepts, des contextes, du sens et non plus simplement des chaînes de caractères. L'autre vise l'exhaustivité en indexant des milliards de pages Web. Aujourd'hui, et en attendant le Web sémantique ainsi que la recherche dite "naturelle" ¹⁰², toutes deux grandes quêtes du monde de l'internet, seul l'apport de communautés d'humains permet d'injecter un peu de sens à une recherche. Mais de son côté, la recherche coopérative comporte bien entendu des risques de subjectivité.

Sur les moteurs de recherche, la coopération peut prendre une autre forme via le "périmétrage" des sites à interroger : la construction d'un moteur sur mesure en quelque sorte. C'est la verticalisation de la recherche.

f. La verticalisation par la personnalisation du périmètre de recherche

L'une des grandes faiblesses des moteurs de recherche web à large échelle (*Google, Exalead, Yahoo!, MSN...*) est le bruit généré par certaines requêtes. Lorsque vous tapez "*puma*", le moteur ne sait pas, à priori, si vous vous intéressez aux vêtements de sport, à l'animal, ou aux hélicoptères. Or, si vous ne réalisez des recherches que dans l'aérospatiale, vous devrez souvent désambiguïser par vos mots clés, ce qui engendre une perte de temps voire de résultats si vos requêtes deviennent trop restrictives.

Certains moteurs ou métamoteurs *clustérisent* (catégorisent) les résultats en les rassemblant par groupes (grappes) dans de petits dossiers. Cette démarche avait été initiée par le moteur *Northern Light* en 1998. Ainsi, tous les résultats relatifs aux chaussures *Puma* sont regroupés dans une même "valise", ceux évoquant l'animal dans une autre. Mais cette solution n'est pas complètement satisfaisante car elle s'effectue à posteriori, une fois votre requête lancée. Le bruit n'est pas éliminé mais simplement circonscrit dans des sortes de catégories "poubelle".

Pour bénéficier d'une démarche à priori, il convient d'utiliser des outils de recherche verticaux c'est-à-dire spécialisés dans un secteur d'activité, une fonction, un type d'information ou de médias.

Face aux problèmes de pertinence des moteurs généralistes, de plus en plus de moteurs verticaux voient le jour. Ainsi, le seul domaine médical a vu apparaître ces 20 derniers mois les moteurs *SearchMedica*, *Medstory*, *MedWorm* ou encore *Medgle*¹⁰³.

Mais si votre secteur de prédilection n'est pas couvert par un moteur spécialisé ?

C'est là qu'intervient un autre type de moteur collaboratif à qui l'on peut indiquer le périmètre des sites sur lesquels effectuer la recherche. Sur les moteurs classiques, les algorithmes peuvent placer en priorité des résultats sans intérêt pour vous. A force de navigation et de veille dans votre secteur, vous connaissez certainement 10, 20, 100 sites que vous jugez incontournables et indispensables pour vos recherches.

2 solutions techniques existent :

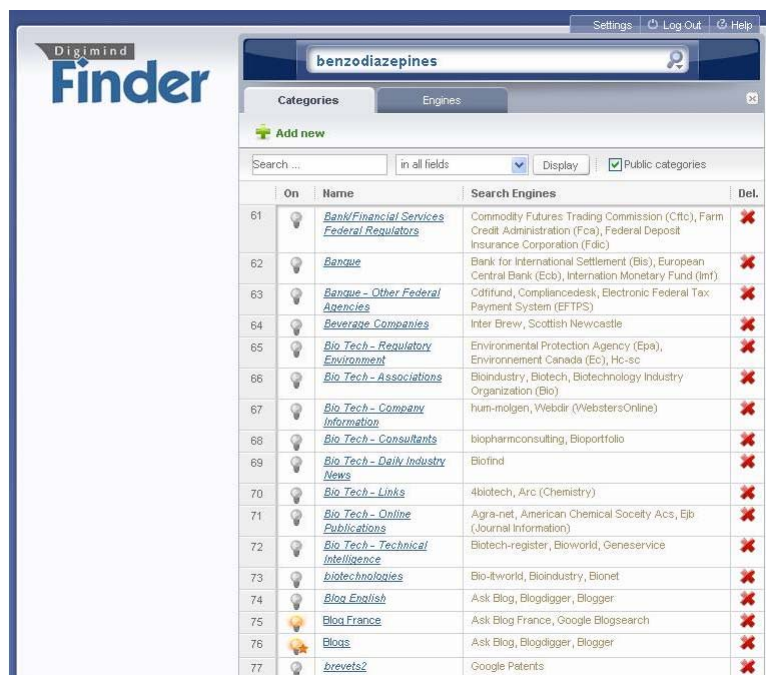
- les outils à périmètre de recherche personnalisable (adaptable scope)

Ils utilisent les index des moteurs généralistes (*Yahoo!* ou *Google*) et restreignent le scope de leur recherche aux pages des sites que vous avez choisis. Ainsi des outils comme *Rollyo*, *Swicki*, *Yahoo! Search Builder*, *Google Custom Search*¹⁰⁴ vous demandent d'indiquer une liste de sites sur lesquels vous allez limiter votre recherche. Lors d'une requête, les résultats proviendront uniquement des sites de votre choix. Pratique.

Mais ces outils présentent une limite de taille : ils ne prennent en compte que les pages indexées par *Google* ou *Yahoo!*. Conséquence : si la réponse pertinente à votre requête figure au sein d'une page non indexée par le moteur généraliste, vous n'obtenez pas de résultats.

De plus, en n'interrogeant que l'index d'un ou 2 moteurs, vous manquerez obligatoirement de nombreuses informations puisque le nombre de résultats communs entre les moteurs diminue constamment¹⁰⁵.

Les pages non indexées par les moteurs généralistes constituent une part du web invisible¹⁰⁶ et sont plus nombreuses que l'on ne croit. Ainsi, si j'interroge le moteur interne du site de la prestigieuse revue scientifique *The Lancet*¹⁰⁷ avec la requête "Diabetes", j'obtiens plus de 2800 résultats. Si j'effectue cette même requête sur *Google* en restreignant ma recherche au site *The Lancet* (via la syntaxe `site:http://www.thelancet.com/`), j'obtiens...1 résultat; Sur *Yahoo! Search*, guère mieux : 9 résultats. Pourquoi ? Parce que les technologies de ces moteurs ne permettent pas d'indexer correctement tous les sites.



▲ La constitution de bouquets de moteurs spécialisés sur *Digimind Finder*.

- les outils à moteurs personnalisables

La solution pour bénéficier des résultats de moteurs internes non indexés par des *Google* ou *Yahoo!* est d'utiliser un outil permettant d'exploiter directement les moteurs des sites qui vous intéressent et ce, sans avoir à passer par les index de moteurs généralistes. C'est le cas du *Finder*¹⁰⁸, mis au point par *Digimind*, un

métamoteur qui permet de *plugger* n'importe quel moteur de son choix et de l'interroger ensuite. Il est donc alors simple de se constituer des bouquets de moteurs spécialisés dans votre secteur d'activité. De plus, cet outil indexe en temps réel les résultats et les documents retournés par les moteurs interrogés. Cela permet d'affiner ensuite vos résultats en interrogeant cet index au contenu spécialisé via une recherche avancée. Avec le *Finder* c'est donc l'utilisateur final qui constitue sa propre sélection de moteurs.

g. L'enrichissement de bases de données façon Web2.0

Nous avons vu que les bookmarks sociaux permettaient de mutualiser les ressources et, pour certains, autour de communautés dédiées comme les scientifiques. Avant l'apparition des applications Web 2.0, ont été constituées de nombreuses bases de données professionnelles. Celles-ci, jusqu'ici, n'offraient pas aux groupes d'utilisateurs la possibilité d'interagir, de commenter et d'échanger. Aussi, de plus en plus de grandes bases de données se dotent d'interfaces collaboratives et ergonomiques, pour laisser plus de place à la contribution.

PubMed

C'est le cas de l'une des plus grandes bases du domaine médical, *PubMed*¹⁰⁹. Cette immense base bibliographique a vu naître deux initiatives¹¹⁰ pour améliorer son interface de recherche :

- Avec *PubMed Interact*, la recherche intègre de l'*Ajax* et permet de pondérer, via des curseurs, les critères de sources, d'ancienneté, de nombre de résultats, de date de publication, de liens avec le mapping du *MeSH* (le thésaurus de la base), de filtrage via des catégories... Vous pouvez ensuite annoter et attribuer une note de pertinence à vos résultats, afficher l'abstract des articles ou leur texte complet, supprimer des résultats ou encore afficher les articles liés. On retrouve donc l'essentiel des fonctionnalités de l'actuel *PubMed* mais avec une convivialité et ergonomie Web 2.0 et donc une plus grande accessibilité.

- *Hubmed* est une autre interface alternative qui ajoute des fonctionnalités à *PubMed* : application de ses propres tags aux résultats de recherche, historique des recherches, clustérisation des résultats, visualisation sous forme graphique, création simplifiée de fils RSS avec les mots de votre requête, suggestion de termes connexes, visualisation de vos mots clés dans l'arborescence *Mesh*. Le tout est développé en *Ajax* notamment, permettant une navigation plus fluide et plus rapide. Parmi toutes ces fonctions, la possibilité de tagger les résultats est très novatrice : on peut ainsi se départir de la taxonomie de la base via sa propre folksonomie.

diabetes
Search Hide Limits

Publication Date: **2004 to 2005**

Journal Subset: **PubMed: full text**

Age Group: **Preschool Child (2-5 years)**

Human/English:

Clear Search | Default Limits | Reload Form
 Slider Presets: [Keyword Search](#) | [Systematic Reviews](#)

Methodology Filter: **Clinical prediction guides (broad)**

MeSH Mapping: **Keywords should be in Title/Abstract**

Citations to Display: **20**

184 citations

Information Box

Total Results: 270927 (10 loaded) [View Results in PubMed](#)
MeSH Terms: "diabetes mellitus"[MeSH Terms] AND "diabetes insipidus"[MeSH Terms]

All Articles High [0] · Low [0] · Unmarked Items [10] · All Items [10]

Auto-append article (A3) · [Show Abstracts](#) · [Hide Abstracts](#) · [Hide Related Articles](#)

1. **Inhaled insulin for diabetes mellitus.**
 Guevara CA
N Engl J Med. 2007 May 17; 356(20):2106-7; author reply 2108
 PMID: 17514779
 Publication Type: [Comment](#) / [Letter](#)

[No Abstract](#) · [Related articles](#) · [PubMed](#) · [Full Text](#) · [Delete](#) · [Top](#)
[Mark relevance as:](#) [High](#) · [Low](#) · [Unmarked](#)
2. **Inhaled insulin for diabetes mellitus.**
 Gross JL
N Engl J Med. 2007 May 17; 356(20):2107; author reply 2108
 PMID: 17514778
 Publication Type: [Comment](#) / [Letter](#)

[No Abstract](#) · [Related articles](#) · [PubMed](#) · [Full Text](#) · [Delete](#) · [Top](#)
[Mark relevance as:](#) [High](#) · [Low](#) · [Unmarked](#)
3. **Inhaled insulin for diabetes mellitus.**

▲ Les nombreux réglages de filtres sur l'interface *PubMed Interact*.

L'USPTO

L'USPTO ¹¹¹ est la base des brevets américains (*United States Patent and Trademark Office*).

Cette vénérable institution a en projet une plateforme de brevets "sociale" : à la manière d'un *Digg* (cf les services d'actualités collaboratives), les internautes pourront commenter les brevets en ligne et les évaluer via un système de notation. Les brevets les mieux notés remonteront alors en haut de classement.

h. Conserver et exploiter l'historique de vos recherches et navigations.

Le chercheur d'information sur le net perd souvent ses nerfs quand il ne retrouve plus ce qu'il a déjà collecté quelques semaines auparavant ou lorsqu'il lui est impossible de se souvenir quelle requête a généré ce fameux résultat remarquable, cette véritable pépite...

Google a résolu partiellement ces problèmes de mémoire. Mais, comme souvent, cela ne va pas sans certains sacrifices.

Via son service *Historique Web*¹¹² (ex *Historique de recherche*), vous pouvez, après enregistrement gratuit :

- mémoriser les requêtes effectuées sur le moteur ainsi que les résultats cliqués,
- enregistrer votre navigation sur les pages web.

Il est alors possible d'effectuer des requêtes sur et à l'intérieur des pages que vous avez visitées dans le passé (sur le texte, les images, les vidéos...).

Google affiche donc désormais vos requêtes, la date et l'heure de la recherche, les résultats cliqués, mais aussi les résultats de recherches qui n'ont pas généré de clics.

Par ailleurs, ce service va encore plus loin dans la personnalisation et l'analyse de votre recherche :

Avec la fonction "*Tendances*" s'affichent vos requêtes les plus courantes, vos sites les plus visités, les résultats que vous avez le plus cliqués.

Plus puissante, la fonction *Éléments intéressants* où Google vous propose des requêtes et des pages web en relation avec vos habitudes de recherche¹¹³.

En outre, si vous activez cet historique, Google l'utilisera aussi pour, à moyen terme, personnaliser vos recherches en fonction de vos centres d'intérêt. Le moteur explique :

"Vos résultats personnalisés seront basés sur vos recherches précédentes, ainsi que les résultats de recherche et les titres à la une sur lesquels vous avez cliqué.

Dans les premiers temps, la différence ne sera peut-être pas notable, mais à mesure que votre historique de recherches s'étoffera, les résultats obtenus seront de plus en plus personnalisés."

D'ailleurs, si vous réalisez de nombreuses recherches sur Google avec une connexion à votre compte Google activée, vous devriez voir s'afficher de temps à autre ce type de message : *"Nouveau : activez votre historique de recherche pour personnaliser vos résultats..."*

Ces résultats personnalisés, apprenant selon vos intérêts, permettent de répondre en partie à la problématique de désambiguïsation lors de requêtes sur des homonymes. Ce sont les exemples classiques de l'avocat (fruit ou juriste ?), de "jaguar" (voiture ou animal ?) ou de "puma" (articles de sports, animal ou hélicoptère ?)

Côté vie privée, tout votre surf "part" sur les serveurs de Google. Aussi, vous pouvez (devez) mettre en pause l'enregistrement de votre navigation lorsque cela vous apparaît adéquat.

The screenshot shows the Google Web History interface. At the top, there is the Google logo, a search bar, and buttons for "Search History" and "Search the Web". Below this is the "Web History for" section. On the left, there is a sidebar with navigation options: "All History", "Web", "Images", "News", "Products", "Sponsored Links", "Video", "Maps", "Pause", "Remove items", "Trends", "Interesting Items", and "Bookmarks". The main content area displays search history for May 7, 2007 (cont.), May 6, 2007, Mar 25, 2007, and Mar 20, 2007. Each entry includes a timestamp, the search query, and the number of results viewed. Some entries also show a star icon and the source website.

Date	Time	Search Query	Results Viewed	Source	
May 7, 2007 (cont.)	9:50am	telecoms OR telecommunications			
	9:50am	telecom			
May 6, 2007	4:29pm	usa	1	usanetwork.com	
	4:28pm	fraises	1	technorati.com	
	4:27pm	charbon	1	wikipedia.org	
	4:10pm	glace vanille	1	glaces.org	
	4:08pm	meteo	1	meteoconsult.com	
	4:08pm	meteo	1	meteo.ch	
	Mar 25, 2007	6:20pm	my yahoo	1	yahoo.com
		6:13pm	levers de soleil		
	Mar 20, 2007	4:27pm	journée internationale de la Francophonie		

▲ L'historique de vos requêtes et de vos résultats sur Google.

i. Les "mash-up"

Les "mash-up" sont certainement les nouveautés engendrées par les technologies dites Web 2.0 les plus spectaculaires. Mixage de plusieurs applications provenant de plusieurs sites (souvent deux), ils permettent de créer un service nouveau et ce, à partir de la mise en relation voire du croisement des données des applications d'origines. En français, Wikipedia parle d'"applications composites"¹¹⁴. Ainsi, en mixant la liste des sociétés où il fait bon travailler établie par le magazine *Fortune* et une carte des USA issue de *Google Maps*, on obtient une carte qui localise toutes ces entreprises. De même, si on mélange les données d'un site de météo à une carte de France, le résultat est une carte des températures et des nébulosités en quasi temps réel.¹¹⁵

Concrètement, comment cela fonctionne-t-il ? Chacune des applications communique et échange des informations via des API (voir glossaire). Ainsi *Yahoo! Maps* ou *Google Maps* ont ouvert leurs interfaces de programmation à tous les développeurs.

Les exemples sont très nombreux car les *Mashups* ont rapidement connu un vif succès auprès des développeurs et internautes avertis qui parviennent à maîtriser le code et les API. Les *Mashups* de cartographie ne sont pas les seules possibilités mais ce sont les plus fréquents, tant la géolocalisation est attrayante pour les professionnels comme pour l'internaute novice. Nous allons explorer quelques mashups pour mieux comprendre leur intérêt.

Les mashups cartographiques les plus courants mélangent des cartes et des données provenant de services sociaux comme *Flickr*, le service de partage de photo de *Yahoo! Inc.* C'est le cas de *Tagmaps*¹¹⁶ qui affiche en temps réel sur une carte les tags associés aux photos de lieux mise en ligne sur *Flickr*. Le Mashup *Flickrvision*¹¹⁷ fait encore plus "simple" : il affiche les photos à l'emplacement de la ville d'origine de l'internaute qui a posté l'image, là aussi en temps réel.

Données statiques et temps réel

Au-delà de certains mixages un peu gadget, certains *Mashups* sont vraiment intéressants pour la surveillance de données ou la collecte d'informations. En effet, on peut trouver deux grands types de visualisation : celle de données statiques (une liste d'adresse, d'entreprises géolocalisées) et celle d'informations évoluant en temps réel. C'est le cas par exemple du Mashup *SpaceCraft Tracking*¹¹⁸ qui permet

de visualiser la trajectoire, sur le globe terrestre, de la navette spatiale ou du télescope *Hubble*.

Souvent, le Mashup exploite des données de répertoires, d'ouvrages, de services publics pour les rendre plus compréhensives, attrayantes, bref plus accessibles ¹¹⁹ :

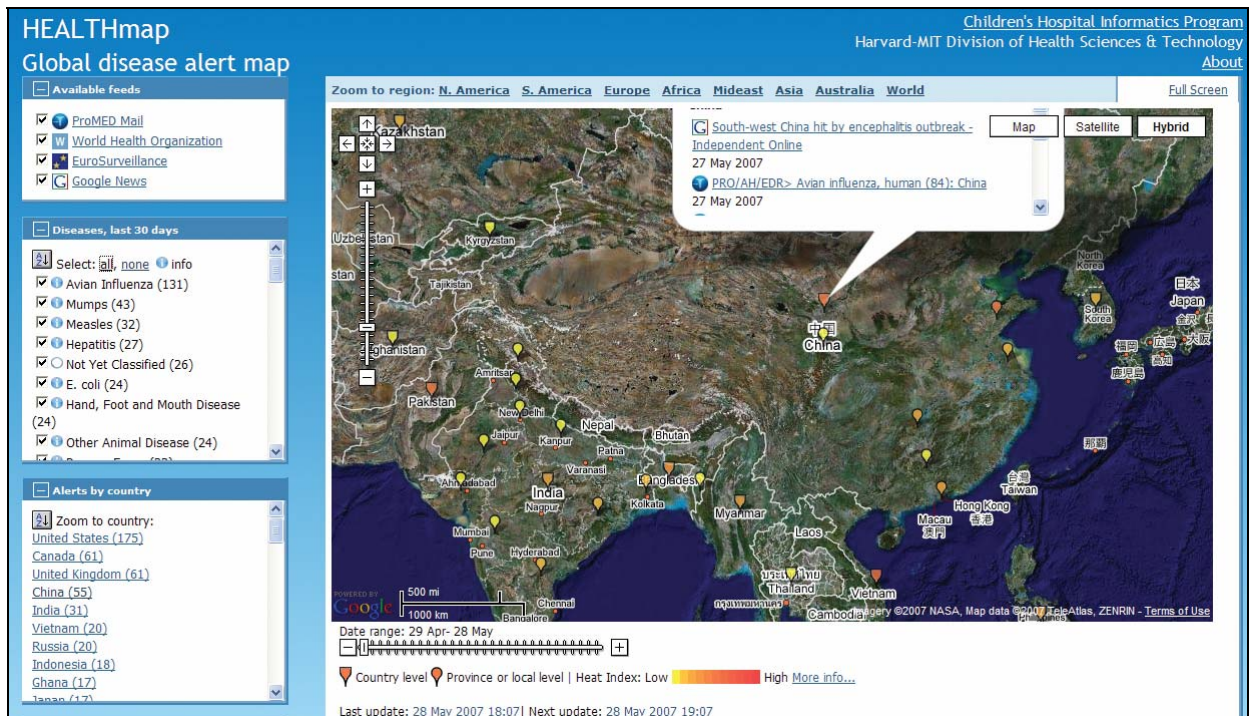
- *HealthMap* rassemble des données de santé publique et des actualités issues de plusieurs sources (*Google News*, *OMS*, *European Centre for Disease Prevention and Control*...) afin de réaliser une carte des maladies infectieuses dans le monde. Cette carte se met à jour au fur et à mesure de l'actualisation des informations.

L'interactivité permet de zoomer sur un pays et de sélectionner des pathologies en particulier.

- *BioDiesel Station Maps* affiche sur une carte les stations services américaines qui délivrent des biocarburants.

- *Kartametro* superpose le réseau du métro moscovite à la carte de la ville. On y visualise aussi les correspondances avec les bus.

- *Wikimapia* mixe les informations relatives à des lieux de *Wikipedia* à une carte *Google Maps*. Résultat : la plupart des noms de villes, sites ou monuments sont cliquables et pointent vers les articles de l'encyclopédie.



▲ Le Mashup *HealthMap* : ici, des actualités sur les maladies infectieuses en Asie, géolocalisées.

-*Global Incident Maps* affiche sur un planisphère les événements liés au terrorisme ou les risques à partir des informations de sites spécialisées mises à jour en continu.

-Enfin, un parfait exemple de visualisation et de surveillance d'événements en temps réel : *DartMaps* indique l'emplacement des trains de banlieue de Dublin.

Mais certains Mashup, en minorité certes, n'exploitent pas de données cartographiques. C'est le cas de *Delexa.org*¹²⁰, un outil qui "mixe" les bookmarks sociaux de *Del.icio.us* aux données statistiques d'*Alexa*. *Delexa* va en effet afficher les tags de *Del.icio.us* associés aux statistiques de fréquentation des 50 000 premiers sites web U.S, fournis par *Alexa* (à partir des données transmises par les barres d'outils *Alexa* installées par les internautes). Un tel outil est intéressant, notamment pour se familiariser avec les tags et *Del.icio.us* mais surtout pour découvrir quels sont les sites d'influence et populaires dans un secteur. On peut également observer la corrélation, ou non, entre nombre de tags et données de fréquentation *Alexa*. Côté moteur de recherche, les mashups *Molu* et *Gnosh* fournissent des métamoteurs réalisés à partir du mixage de dizaines d'outils de recherche¹²¹.

Des ressources

Comment trouver des mashups ? Comme pour de nombreuses informations sur le web, des annuaires spécialisés¹²² sont apparus. Ils répertorient aussi bien des API à destination des développeurs que des mashups. *WebMashup*, auto-proclamé *Open Directory for Mashups et Web 2.0 APIs* en liste des centaines. *ProgrammableWeb*, l'un de sites de référence dans le domaine en propose plus de 1800 sous la forme d'un répertoire et également d'une matrice, permettant de visualiser clairement le "croisement" de deux applications ou plus. *Mashup Feed* liste aussi des mixages populaires.

Des blogs sont également centrés sur ces pratiques : *Google Maps Mania*¹²³ recense de nombreux Mashups cartographiques.

Ces Mashups sont un beau symbole de ce web nouvelle génération : on ne lit plus seulement de l'information statique de manière passive. Désormais, on est à la fois fournisseur et spectateur en temps réel des contributions de communautés. Le tout

grâce à des interfaces réalisées par des développeurs à partir de code ouverts de grandes sociétés. Le succès est notamment provoqué par la possibilité de mixage entre des données du monde réel et des informations ou fichiers postés sur le net. Les internautes peuvent donc maintenant se rassembler par centres d'intérêt mais également par proximité géographique : *Je recherche un spécialiste de l'information sur les biocarburants, situés à moins de 10 km de chez moi.*

Par ailleurs, on peut imaginer, à terme, des applications mélangeant les données de centaines de fichiers et bases de données. On en est certes aux prémises mais le mashup *Alkemis* ¹²⁴ en est déjà un bon exemple : il affiche, sur une carte de New York, des données provenant d'*A9 OpenSearch, Del.icio.us, Flickr, Google Maps et Yahoo Traffic*. Résultat : la possibilité de visualiser en une seule interface un état du trafic routier via des caméras de surveillance, la météo, les actualités locales ou les liens *Del.icio.us* en rapport avec la ville....

III. Diffuser

Avec l'apparition des applications collaboratives et des plateformes de blogs, la facilité de publication sur le web s'est accrue : Tout individu, même sans connaissances techniques, peut publier et diffuser son propre contenu sur un blog, un wiki et via des fils RSS.

a. Les Blogs et le flux RSS

Il suffit de 5 minutes pour s'inscrire à une plateforme de blogs et commencer à écrire. Ces plateformes proposent des services de bases gratuits et des fonctionnalités plus avancées nécessitant un abonnement. Ces services qui permettent de créer, maintenir et héberger votre blog sont nombreux. Parmi les prestataires de qualité, on peut citer *Blogger* (qui appartient à Google), *TypePad* et *Vox* (de la société *SixApart*), *Overblog*, *Viablog*, *Haut et Fort* ou *BlogSpirit*¹²⁵. Bien d'autres sociétés proposent des services de création et d'hébergement¹²⁶. Des comparatifs sont disponibles en ligne¹²⁷.

Nous l'avons vu, le format RSS permet de diffuser très facilement le contenu de son blog : utilisant le XML (un html épuré), langage natif du blog, chaque carnet web comporte un fils RSS permettant aux visiteurs de s'abonner à vote blog et de recevoir instantanément toute nouvelle publication : actualités, billets, podcast ou vidéos. Par ailleurs, des services comme *Feedburner*¹²⁸ permettent de rendre compatibles vos flux RSS avec la majorité des lecteurs RSS, outils dont nous reparlerons plus loin.

La création de sites internet "classiques" s'est elle aussi considérablement simplifiée : un outil tel que *Google Page Creator*¹²⁹ permet de créer un site simple en 1 heure. Des sites plus sophistiqués sont également abordables sans avoir à connaître une ligne de programmation de html grâce au perfectionnement des éditeurs dits WYSIWYG : ce que vous voyez quand vous construisez votre page est ce que vous obtiendrez en ligne.

Les outils Web 2.0 permettent à des spécialistes, journalistes ou salariés de diffuser leur contenu. Il y a encore 2 ans, ils étaient pour la plupart écartés de la publication web faute de bagage technique.

Depuis le développement des blogs à partir de l'automne 2001, il est possible pour tout internaute de diffuser sa propre information. Au-delà des blogs d'adolescents et autres "carnets intimes", le web regorge de blogs à tendance professionnelle : corporate, produits, marketing, experts, politiques, ressources humaines. C'est autant de sources potentielles qu'il faut exploiter afin de capter des sources primaires de qualité. Le volume des blogs publiés et maintenus chaque jour est lui aussi en constante augmentation.

Le White Paper "*Blogs et RSS : des outils pour la veille stratégique*" développe les aspects intéressants des blogs pour le professionnel de l'information¹³⁰.

b. Les Wikis

Le wiki est finalement différent du blog de par sa nature et son utilisation. Selon *Wikipedia*, le Wiki est un "site Web dynamique permettant à tout individu d'en modifier les pages à volonté. Il permet de communiquer et diffuser des informations rapidement, de structurer cette information pour permettre d'y naviguer commodément". (Wiki vient du terme hawaïen wiki wiki = "rapide" ou "informel").

Là où le blog sera adapté à la diffusion d'actualité et de contenu se renouvelant fréquemment, le wiki sera davantage dédié à du texte plus statique mais nécessitant ou pouvant s'enrichir de la collaboration de plusieurs personnes. Ainsi, les articles d'une encyclopédie ou les fiches correspondants aux différentes phases d'un projet impliquant de nombreux collaborateurs pourront être réalisées sous forme de Wiki.

Sur un blog, l'affichage du contenu est anté-chronologique (le plus récent en haut), sur un wiki, il est structuré en rubriques, l'architecture comme le contenu pouvant être modifiés à tout moment par les participants. Ceux-ci peuvent bénéficier de droits d'accès et de modification hiérarchisés.

Les versions antérieures ainsi que les modifications sont accessibles via une rubrique "*Historique*". Sur un blog, le visiteur peut collaborer mais uniquement via ses commentaires. Sur un wiki, tout collaborateur autorisé peut modifier le texte comme le titre et l'organisation des rubriques.

Par ailleurs, la prise en main d'un wiki et de ses syntaxes d'enrichissement et modification demande un peu de pratique et n'est pas aussi simple que le maniement d'un blog. Le wiki est un des outils idéaux pour la gestion de projet. Certaines entreprises les utilisent en tant que tel (parfois en parallèle avec les blogs) pour éviter des échanges, des recherches et un archivage via les mails fastidieux.

Géré via des droits d'accès gradués, il peut alors se révéler comme un outil très simple à manier et à s'approprier. Ces wikis seront évidemment réservés à l'entreprise en accès privé.

Exemples de wikis publics

> Wikipedia¹³¹

Le wiki le plus connu est Wikipedia, l'encyclopédie collaborative gratuite dont le contenu est créé et maintenu par des milliers d'internautes bénévoles. Chacun est libre de créer un article, modifier ceux des autres ou les commenter. En mai 2007, l'encyclopédie proposait plus de 7 millions d'articles dans 192 langues et dialectes dont plus de 500 000 en français. Wikipedia jouit d'une popularité et d'un succès très important. Google indexe de nombreux articles de Wikipedia à tel point qu'il est rare d'effectuer une recherche sur le moteur sans voir s'afficher une définition issue de l'encyclopédie. Selon la société de mesure d'audience Comscore, en janvier 2007, Wikipedia était le 9ème site US le plus visité (derrière les sites Yahoo!, premiers, puis les sites Time Warner, Microsoft, Google, eBay, Fox Interactive, Amazon et Ask). 6 groupes éditant un moteur généraliste (Google, Ask, Microsoft, Yahoo!, Amazon et AOL) occupent le top10 US. Mais c'est la première fois qu'un site collaboratif trône les premières places des sites les plus visités.

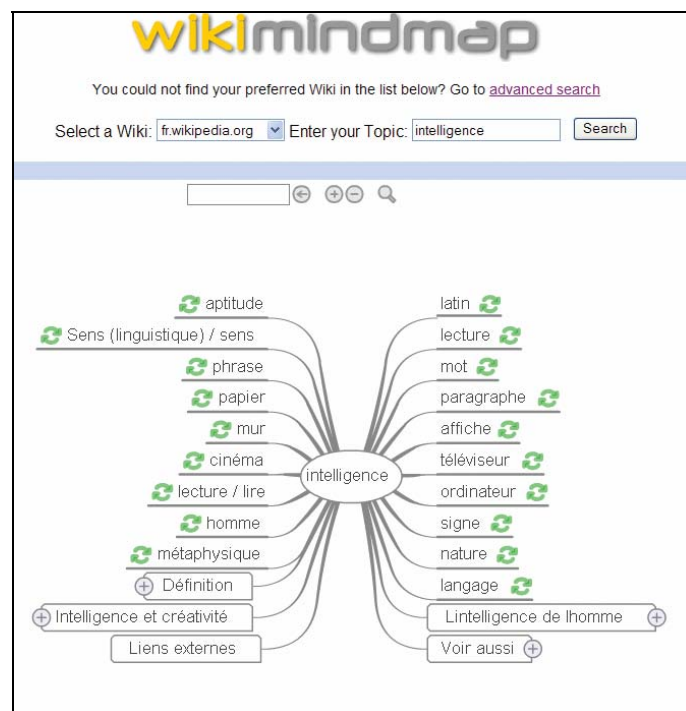
Wikipedia, comme tout service leader, est très critiqué. Notamment sur la possibilité pour tout un chacun de modifier et de créer des articles pouvant entraîner des risques de diffamation ou de "vandalisation"¹³². Même si ces risques sont réels, nous le verrons dans la partie consacrée aux limites du Web 2.0, il ne faut pas pour autant se focaliser sur ces problèmes.

Par ailleurs, on a souvent mis en cause la qualité des articles : du contenu émanant d'internautes pas forcément qualifiés ou experts et donc par nature douteux. Or, une étude commandée par la Revue *Nature* à remis en cause ce type de ressenti : des experts ont comparé 42 documents publiés sur Wikipedia et sur la célèbre *Encyclopaedia Britannica*. Résultat : la qualité des deux sources est assez proche¹³³. Seulement 8 erreurs sérieuses comme des mauvaises interprétations de concepts ont été repérées dans les articles analysés, 4 pour chaque encyclopédie. Des erreurs factuelles, omissions ou fausses descriptions ont également été détectées : 162 dans Wikipedia et 123 dans l'*Encyclopaedia Britannica*.

Le succès et l'importance de Wikipedia n'ont pas échappé aux moteurs de recherche qui sont plusieurs à s'être spécialisés dans la recherche sur l'encyclopédie. Ainsi WikiSeek¹³⁴ est un moteur lancé début 2007 qui recherche sur le contenu de Wikipedia. Il indexe le contenu en crawlant les pages en profondeur et "aspire" également le contenu des liens présents sur les pages de l'encyclopédie.

Plus élaboré, le moteur français *Exalead*¹³⁵ enrichit considérablement la recherche dans l'encyclopédie en ligne. Cette fonctionnalité est accessible en cliquant sur l'onglet *Wikipedia*. Il est évidemment possible de rechercher des articles par mots clés mais aussi d'affiner par langue, via le "nuage de tags" dynamique (les étiquettes correspondent à des emplacements géographiques, personnes célèbres, noms d'organisations, termes liés à votre requête ou catégories de Wikipedia).

*WikiMindmap*¹³⁶ recherche sur Wikipedia et affiche les résultats sous forme de MindMap (cartes heuristiques). Il est alors possible de naviguer dans l'arborescence des termes.



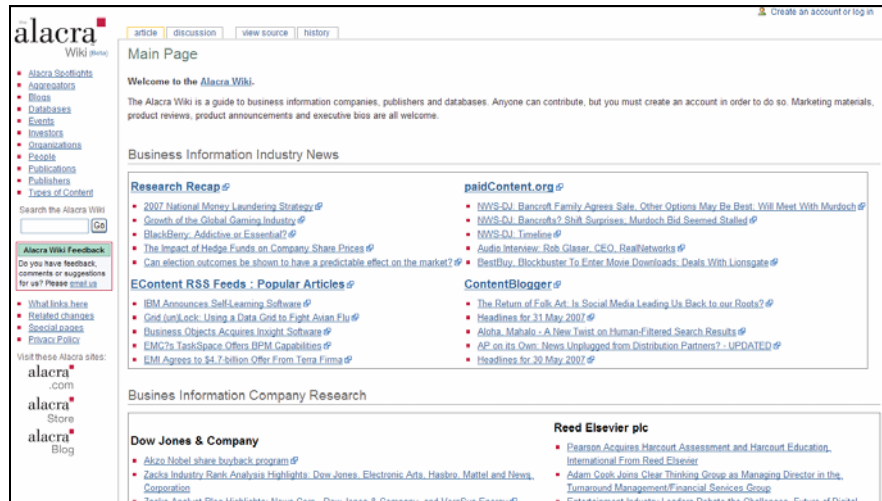
▲ WikiMindMap

> Citizendium¹³⁷

Lancé fin 2006, *Citizendium* est une initiative de Larry Sanger, co-fondateur de Wikipedia avec Jimmy Wales. Citizendium est également une encyclopédie. Mais, jugeant la communauté Wikipedia, sensée s'auto-réguler, dysfonctionnelle, Larry Singer intègre dans Citizendium des procédures de sélection des articles par des experts et refuse aux contributeurs la possibilité de l'anonymat, misant ainsi sur la responsabilisation. Le collège d'experts a le dernier mot concernant la vérification des articles, le but étant d'obtenir une encyclopédie contenant le moins d'erreurs possible. Citizendium est structurée en grandes catégories académiques correspondant à des groupes de travail : Sciences naturelles, Sciences Humaines, Arts, Sciences Humaines, Divertissement, Arts appliqués et Sciences.

> Alacrawiki ¹³⁸

AlacraWiki est un guide dédié aux sociétés du secteur de l'information, éditeurs et bases de données. La rubrique "Alacra Industry Spotlights" est particulièrement remarquable : elle fournit des analyses et commentaires sur les sites d'intérêt pour une trentaine de secteurs industriels. Le wiki offre des informations sur les agrégateurs de contenu, les bases de données et les meilleurs blogs dédiés au secteur de l'information on-line. Alacra propose aussi un blog et un fils RSS.



▲ Alacra Wiki

> "Les fils RSS" ¹³⁹

Sur cette page de wiki, vous trouverez à peu près tout sur le RSS :

La norme, comment utiliser les fils RSS, des articles, une sélection d'agrégateurs, mixer des fils RSS entre eux, mesurer l'activité des fils RSS, créer un fil RSS....

> Wikidweb ¹⁴⁰

Wikidweb est un Wiki qui propose un répertoire de sites web (et non de wikis). Aussi, vous pouvez y participer selon les fonctionnalités courantes du format wiki : vous avez ainsi la possibilité d'y ajouter vos propres sites, mais également de modifier les entrées des autres internautes (urls, texte) ou de les commenter.

IV. Surveiller

Nous l'avons vu dans la partie "Recherche de blogs et de flux RSS" : les flux RSS sont présents sur de nombreux formats du net et vont permettre de surveiller les nouveautés apparaissant sur les blogs, les sites corporate, les sites de presse, dans les résultats de moteurs de recherche d'actualités, de blogs, web, dans les newsgroups, sur certains forums web, sur certaines listes de diffusion, les bookmarks sociaux, les sites de partage de photos et de vidéos,....bref sur de très nombreux sites classiques et ressources collaboratives de type 2.0.

Ces flux RSS peuvent être intégrés à votre process de veille en deux clics, via des outils simples gratuits, ou payants mais plus puissants.

a. La détection du flux RSS

Pour pouvoir mettre en surveillance le flux RSS (RSS Feed en anglais), il faut d'abord le trouver. Or, il n'est pas toujours aisé de mettre la main sur ces petits logos oranges ou les liens "RSS", "Abonnez-vous", "Syndiquez ce flux" qui pointent vers les flux RSS parfois cachés dans les méandres des pages web.



▲ Des icônes signalant la présence de flux RSS.

Certains navigateurs permettent de détecter automatiquement les flux lorsqu'ils sont présents sur une page web. Une icône vous avertira alors et, dans la plupart des cas, affichera l'adresse du ou des flux disponibles sur la page. Plusieurs formats de flux peuvent coexister : les différentes versions du RSS (0.90, 0.91, 0.92, 1.0, 2.0) et le format ATOM, concurrent mais offrant les mêmes fonctionnalités

Internet Explorer 7, Opera, Firefox permettent de détecter les flux. C'est aussi le cas de *Maxthon*, une puissante interface de navigation qui vient se greffer sur l'architecture d'*Internet Explorer* ¹⁴¹.

b. La mise en surveillance

Quelque soit l'aspect de la page qui contient votre flux RSS (.xml, .rdf, interface *Feedburner*), il convient de copier et coller l'adresse (url) dans l'outil qui va vous servir à lire ce flux et à être alerté dès lors qu'une nouveauté apparaît.

4 grands types d'outils peuvent être utilisés (du plus simple au plus sophistiqué) :

- Les navigateurs web.

Les navigateurs cités précédemment, outre la détection du flux RSS, permettent de s'y abonner et de les lire. Les flux s'intègrent alors dans un menu latéral ou dans la page principale. C'est simple mais dépourvu de fonctions avancées de tri ou d'organisation.

- Les lecteurs (ou agrégateurs) de flux RSS.

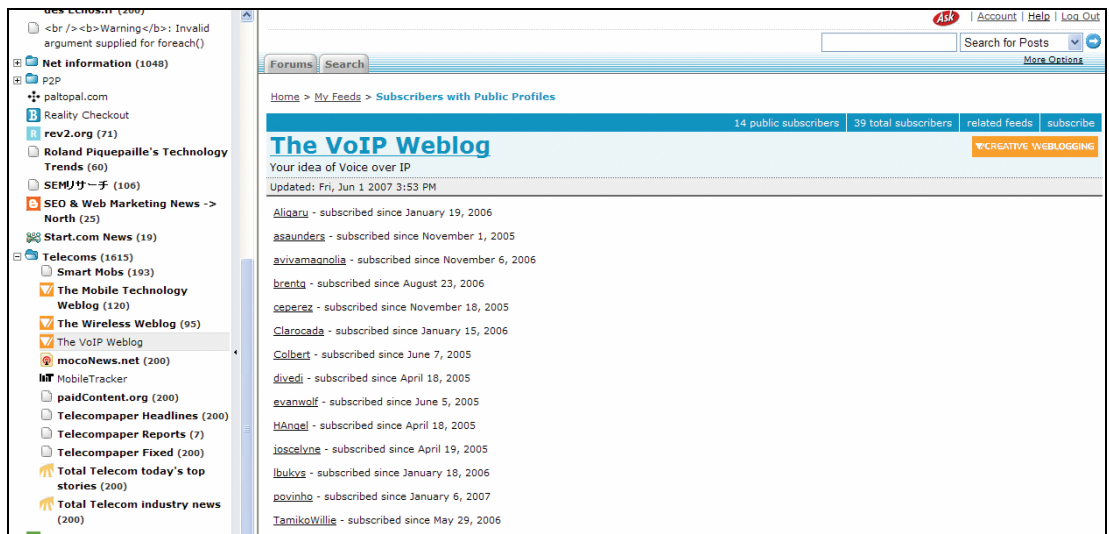
Ce sont les outils dédiés aux fils RSS. Comme de nombreuses applications, ces lecteurs sont soit disponibles en ligne, soit à installer localement sur votre PC.

- Exemples de lecteurs RSS en ligne : *Bloglines*, *Pluck*, *Newsgator*

- Exemples de lecteurs RSS installés : *RSS Reader*, *Amphetadesk*, *Newsgator* (s'intègre à Outlook), *FeedDemon*.

Le lecteur *Bloglines* exploite de manière astucieuse la logique communautaire : lorsque vous êtes abonné à un flux RSS, vous pouvez visualiser les autres abonnés (ceux qui ont accepté de rendre leur abonnement public). Un clic sur leur nom affiche tous les fils RSS auxquels ils ont souscrits : un moyen rapide de découvrir d'autres ressources pertinentes collectées par des internautes proches de vos centres d'intérêt.

Les lecteurs RSS sont très nombreux. Pour mieux les distinguer, certains blogs ou sites proposent des comparatifs¹⁴².



▲ La communauté sur *Bloglines* : les abonnés au flux RSS du blog "Voip Weblog"

- Les pages personnalisables

Ce sont des pages web généralement organisées en onglets et boîtes, permettant d'organiser et de ranger ses flux par catégories, et ce, dans un même espace. Ces services sont ergonomiques et idéaux pour lire d'un coup d'oeil le maximum de flux. Les plus connus sont *NetVibes*, *PageFlakes* ou *Webwag*.

Outre la lecture de flux RSS, ces pages offrent la possibilité d'intégrer des services de météo, mail, bloc-notes et des flux provenant de service sociaux comme les bookmarks ou les réseaux d'amis.

Certains blogs fournissent des listes et comparatifs de ces services¹⁴³.

- Les progiciels de veille intégrés

Face à l'abondance des flux RSS et blogs, des applications de veille intégrée comme *Digimind Evolution*¹⁴⁴ donnent la possibilité de filtrer les RSS via l'application de requêtes différentes sur chaque fils RSS.

On peut ainsi sélectionner les informations à remonter via des requête simples ou plus complexes (booléens, parenthèses, troncatures et opérateurs de proximité), à longueur illimitée. En outre, il est possible de surveiller jusqu'à 10 000 flux RSS quotidiennement.

▪ Les boutons intégrés

Certains sites vont vous simplifier la tâche : ils affichent sur leurs pages, à coté de leurs flux, des boutons qui permettent d'ajouter directement le flux à votre lecteur.

Ainsi, le service d'actualités *Wikio* ou le moteur *Ask Blogs & Flux* proposent tous deux des boutons vers *Bloglines*, *MyYahoo!* *Google Reader*, *Newsgator* et *Ask Blog*.

▪ Les fichiers OPML

L'OPML (*Outline Processor Markup Language*) est un format développé en XML qui permet l'échange d'informations structurées entre applications fonctionnant sur des OS et environnement différents. Exemple : les abonnements d'un internaute à un ensemble de flux RSS entre un navigateur Firefox et Internet Explorer, de MacOS à Windows XP.

Des agrégateurs RSS comme *Bloglines*, *NewsGator* ou *Rojo* offrent la possibilité d'exporter l'ensemble de vos flux RSS favoris ou ceux d'autres abonnés au format OPML pour les importer ensuite dans votre agrégateur.

*Share your OPML*¹⁴⁵ est, comme son nom l'indique, un portail qui vous permet de partager vos flux RSS au format OPML. Ce service liste des centaines de fichiers OPML.

c. Pourquoi une veille via le RSS ?

Faire une veille via les fils RSS présente des intérêts notables :

- L'hyper-segmentation : le choix des fils RSS rend possible une sélection fine de l'information. Ainsi, sur le site du *Figaro*, on peut choisir l'actualité internationale, nationale, les sciences et médecine, évitant ainsi d'être submergé par l'information superflue. Sur *La Tribune*, on peut sélectionner un flux par société.

Certains blogs proposent même des flux RSS permettant aux choix de surveiller les commentaires ou uniquement les billets.

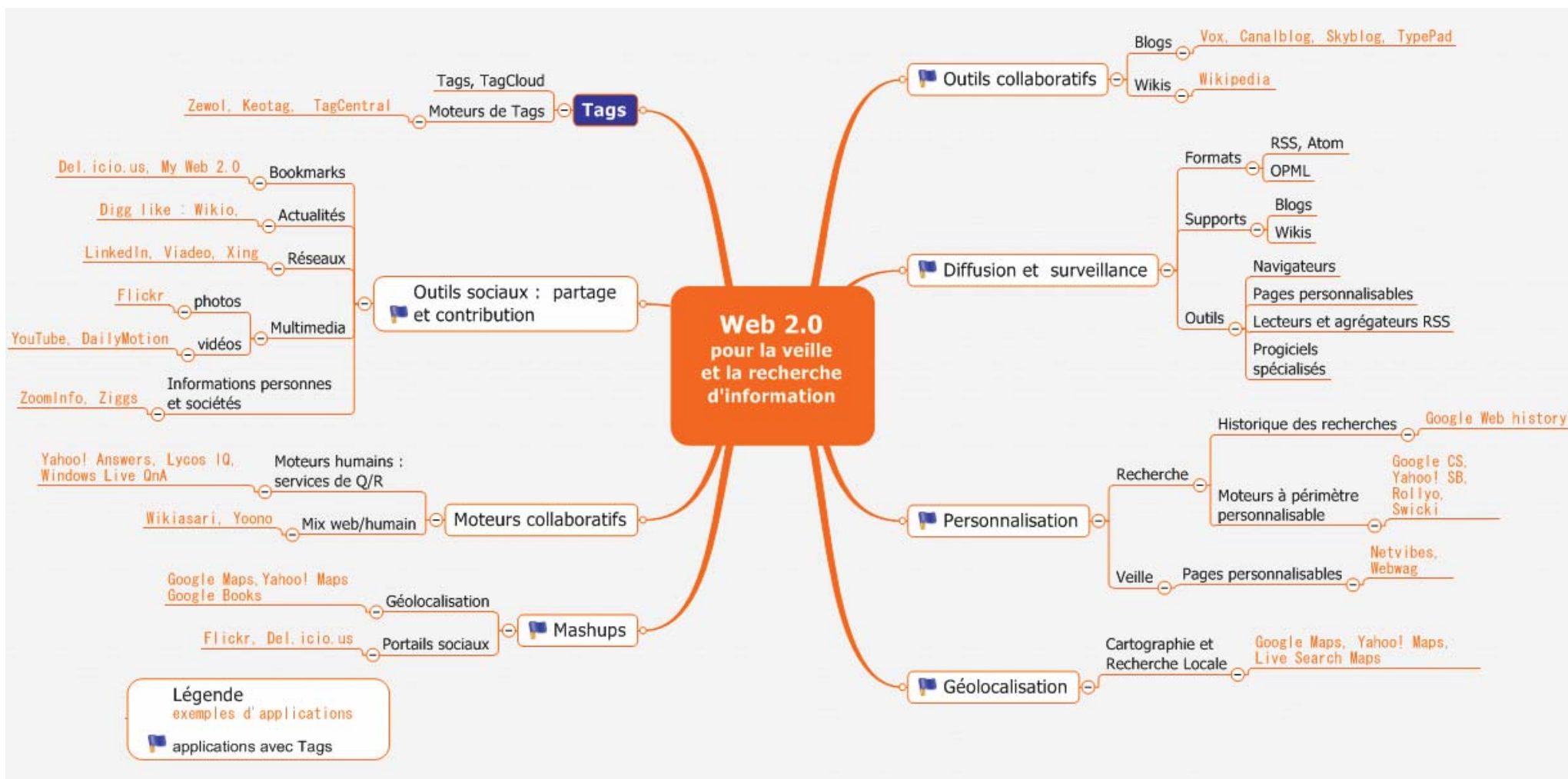
- Un gain de temps : inutile de se connecter au site, au blog ou au moteur de recherche pour visualiser ses nouveautés.

- L'anonymat : l'abonnement à un flux RSS, contrairement à une newsletter ou à une liste de diffusion, est anonyme.

- La consultation aisée : vos fils RSS, via une lecture sur un agrégateur en ligne sont accessibles de n'importe quel PC connecté à l'internet, où que vous soyez.

V. Carte des outils et ressources Web 2.0

> [Page suivante](#)



▲ Carte des outils et ressources Web 2.0 - DIGIMIND Services. 2007.

EXEMPLE D'UNE RECHERCHE ET VEILLE VIA LE WEB 2.0

Cas d'école : Nous avons à mener une recherche et une veille sur le développement des *smartphones* ou téléphones multimédias haut de gamme. Des sites de type 2.0 pourront se révéler complémentaires de sites classiques (sites d'entreprises, portails, associations professionnelles...) pour détecter les tendances et sujets émergents.

Une première recherche via l'encyclopédie *Wikipedia* permettra d'obtenir une première définition du sujet, et d'identifier les principaux mots clés, ainsi que des liens vers des sites pertinents.

Etape suivante : nous pourrons élargir notre recherche en identifiant des sites similaires aux sites trouvés précédemment, via les suggestions fournies par des moteurs collaboratifs tels que *Yoono* ou *Similicious*.

Une recherche sur le tag "Smartphone" dans des sites de social bookmarking, comme *Del.icio.us*, nous permettra également d'identifier les pages plus ou moins "populaires" taggées par les contributeurs du site, ainsi que d'autres tags en relation avec notre sujet.

Explorons ensuite les avis d'"experts" plus ou moins autorisés que l'on trouvera dans les méandres de la blogosphère, grâce à des moteurs de blogs comme *Technorati*, *Blogpulse* ou *Google Blog Search*.

La recherche pourra s'effectuer via les tags ou le texte intégral des billets. On pourra ensuite parcourir les blogrolls des auteurs les plus pertinents, afin d'identifier d'autres sources.

La lecture des blogs fournira non seulement des avis de spécialistes ou journalistes, mais aussi des avis plus ou moins critiques de consommateurs ou utilisateurs de smartphones...

Il conviendra toutefois de rester vigilant quant à l'effet "pot de miel" de certains blogs, où les contributions des internautes se révèlent être de simples messages publicitaires déguisés ! (voir à ce sujet la mise en garde du site *UFC Que choisir*¹⁴⁶).

Pour recueillir les dernières informations d'actualité, on pourra utiliser un agrégateur tel que *Wikio* ou *Digg*, et créer un flux RSS à partir d'une requête sur les

tags précédemment identifiés. *Wikio* propose des raccourcis pour intégrer ce flux au sein de plusieurs lecteurs RSS du marché (*Netvibes*, *Newsgator*, *Google Reader* etc).

Les réseaux sociaux professionnels permettront également d'identifier des acteurs importants du domaine, dans les sphères universitaires ou de l'entreprise.

La communauté 2.0 s'exprimera également sur les grands portails consacrés aux nouvelles technologies. C'est le cas par exemple du portail ZDNet, qui permet aux utilisateurs de commenter les articles, ou de les ajouter directement à leur profil *Del.icio.us*... Les commentaires ou les trackbacks pointant sur un article pourront constituer un précieux complément d'informations.

Afin de mener une veille régulière sur le sujet, on syndiquera les flux RSS provenant des différents sites ou blogs identifiés. Pourquoi ne pas également créer à partir de ce bouquet de sources personnalisées un moteur sur mesure, conçu grâce à *Rollyo* ou *Yahoo! Search Builder* ?

Les limites du Web 2.0

LA PARTICIPATION

I. Une participation hétérogène

Nous l'avons déjà évoqué : tout le monde ne participe pas de la même manière sur les applications du Web 2.0.

Une minorité participative...

Bill Tancer, de la société de mesure d'audience *Hitwise* révèle dans une étude¹⁴⁷ que sur les sites type Web 2.0, les internautes regardent mais participent peu : Seulement 0,16 % des visiteurs du site *YouTube* viennent y partager leurs vidéos. Les 99,84 % restant ne font que regarder. Idem pour le site de partage de photos de *Yahoo!*, *Flickr* : seulement 0,2 % des visites conduisent à la publication de photos. *Wikipedia* est l'exception : sur l'encyclopédie collaborative gratuite, 4,6 % des visites correspondent à une modification des articles.

Sur ces sites, on est donc encore loin de la règle de la 1/10/89 édictée par Charles Arthur dans *The Guardian*¹⁴⁸. Il estime que sur le net, sur un groupe de 100 personnes, si une crée du contenu, 10 vont "interagir" (commentaires, améliorations) et les 89% restant ne feront que le lire et regarder. Le journaliste l'illustre via des exemples quantifiés avec *YouTube*, *Wikipedia*, ou *Yahoo! Groups*.

"*Most Web 2.0 Users Are Really Just Couch Potatoes*" : Bruce Nussbaum compare même les internautes Web 2.0 aux spectateurs de TV passifs et affalés dans leurs canapés¹⁴⁹.

Mais parallèlement, les sites de type Web 2.0 ont vu leur audience grimper de...668% en 2 ans, révèle l'étude *Hitwise*. Cela représente 12% du trafic web aux USA contre 2% il y a seulement 2 ans.

Les sites de partage de photos sont parmi les plus populaires : *Flickr* bien sûr mais aussi *PhotoBucket* qui représente à lui seul 41% du trafic des sites de photos en ligne.

....d'autant plus que la forme de participation se complexifie.

D'après le cabinet *Forrester*, les internautes sont nombreux sur les actions à faible niveau d'exigence (lire, enregistrer un signet, commenter) mais beaucoup moins sur des participations plus actives (écrire, collaborer).¹⁵⁰

Wikipedia, un faux modèle collaboratif ?

On cite souvent l'encyclopédie Wikipedia en exemple : c'est l'un des emblèmes les plus forts et les plus lus du Web 2.0. Son mode de fonctionnement est caractéristique voire symptomatique du web collaboratif. Mais le process de Wikipedia serait finalement très classique avec une hiérarchie de contributeurs, sorte de "cadres" disposant de nombreux droits écrivant la majorité des articles. C'est l'analyse de Lionel Barbe du CNRS¹⁵¹. Il explique que, si le principe de liberté de publication sur Wikipedia reste valable, il est de moins en moins commun car *"la plupart des entrées basiques, ne prêtant pas à discussion, ont déjà été effectuées". "L'étude des « logs » historiques montre que la création de nouvelles entrées qui correspondent à de nouvelles pages articles dans la Wikipedia Francophone, ne concerne qu'environ 3% à 4% des opérations. Par comparaison, cette proportion tombe à 2% pour la version anglaise de Wikipedia qui contient plus de 1 million d'articles. A l'inverse, la proportion monte entre 8% et 10% pour des Wikipédias ayant moins de 100 000 articles comme la Wikipedia en Lituanien et bien d'autres".*

Selon lui, on assiste à un effet de professionnalisation de l'encyclopédie. *"[...] les plus gros contributeurs, que nous considérerons comme étant les 5% des individus éditant le plus d'articles, éditent entre 85% et 95% du contenu total de l'encyclopédie. Au fur et à mesure, ces gros contributeurs acquièrent des privilèges, comme le statut d'administrateur, qui leur permettent non seulement de participer mais aussi de surveiller et le cas échéant censurer le contenu édité par les autres contributeurs et de participer aux grandes lignes de l'organisation de Wikipedia".*

Cette analyse a toutefois été battue en brèche par Aaron Swartz, dans une petite étude¹⁵² : En mesurant, non pas le nombre d'interventions de chaque contributeur, mais la quantité de texte produit par chacun, la réalité est différente : les contributeurs occasionnels écriraient des textes longs en peu d'interventions tandis que les éditeurs actifs réaliseraient un travail tenant davantage de la maintenance (correction, vérification, suppression) mais sans apporter beaucoup de texte.

II. Une participation d'"amateurs" ?

Tout internaute peut-il s'improviser journaliste, expert, commentateur...? Michel Fingerhut ¹⁵³ voit dans le succès d'une encyclopédie comme Wikipedia une manifestation du "web comme hégémonie de l'amateurisme et comme encouragement à l'extrémisme idéologique et au communautarisme".

Les contenus créés et diffusés sont-ils de qualité ? Le Web 2.0 devient une plateforme qui stimule la créativité et la production de contenu par cela même qu'elle en permet la diffusion (avec indice de notoriété). On assiste à une libération de talents qui peuvent enfin se révéler parallèlement à des productions caractérisées par un certain amateurisme et l'absence de véritable ambition créatrice. Car comme sur l'ensemble de l'internet, sur ce Web 2.0, le pire côtoie le meilleur.

Par ailleurs, toujours en prenant l'exemple de Wikipedia, un certain nombre de phénomènes se révèlent tout de même préoccupants :

Si l'étude réalisée pour le magazine *Nature* révèle que la qualité des articles de l'encyclopédie est presque au niveau de celle de l'*Encyclopaedia Britannica*, certains articles de Wikipedia relèvent de la désinformation ou de la propagande :

- l'article "*Droits de l'Homme en Iran*"¹⁵⁴ a ainsi été très controversé : en effet, cet article très long de plus de 20 pages se présente davantage comme une apologie du régime iranien qu'une analyse argumentée des droits de l'homme dans ce pays. Or, certainement au nom du principe de neutralité qui régit Wikipedia, cet article tendancieux a été classé comme "*article de qualité*"¹⁵⁵.

- le cas *Seigenthaler* ¹⁵⁶ : en mai 2005, un contributeur publie anonymement sur Wikipedia un article sur John Seigenthaler Sr., un écrivain célèbre aux USA. Or ce billet mentionne que l'écrivain a été directement impliqué dans les assassinats de John et Bob Kennedy, en précisant toutefois que rien n'a jamais été prouvé. Cette information est tout simplement fautive. Mais l'article n'est découvert (par un ami de John Seigenthaler) et corrigé que 4 mois plus tard, mettant à mal la crédibilité de l'encyclopédie. A la suite de cette affaire largement relayée par les médias américains, le fondateur de Wikipedia a rigidifié les règles de publication concernant les tout nouveaux contributeurs.

Certes, sur les milliers d'articles mis en ligne par Wikipedia, ces dysfonctionnements restent heureusement minoritaires mais ils révèlent tout de même certaines limites du tout collaboratif et la nécessité de règles de publication empruntées aux médias "classiques".

L'agrégation de "petites choses faciles"

Elles ont tendance à dominer le contenu du Web 2.0 : La facilité de publication, le rythme des blogs qu'il convient d'alimenter fréquemment peut entrer en conflit avec le temps de la recherche et de la réflexion. Souvent, les thèmes traités ne peuvent dès lors être que des sujets du quotidien, des opinions non réfléchies, non argumentées... On peut donc craindre, à l'instar du mode de diffusion sur d'autres médias, qu'intégré à un réseau social, noté, évalué, à la recherche de notoriété et de visibilité, le créateur de contenu ait tout intérêt à aller dans le sens de son audience, produire du divertissant plus à même de lui garantir un succès immédiat.

Par ailleurs, des opinions peuvent-elles se transformer en jugement pertinent par l'effet de leur agrégation ? L'agrégation d'une quantité importante d'avis peut elle vraiment dépasser l'imprécision des avis pris indépendamment. Peut-on sans une harmonisation des critères d'appréciation atteindre un jugement collectif fiable ? Les tendances qui ressortent risquent en effet de se limiter souvent au reflet de l'opinion majoritaire.

Mais l'évocation de ces limites ne doit pas occulter toutes les contributions matures, produites par des internautes soucieux d'apporter un contenu véritablement constructif. Les billets et les commentaires de nombreux blogs politiques français en 2007 en sont une parfaite illustration.

LE WEB 1.0 RESTE NECESSAIRE**Le mail résiste**

Nous l'avons vu au cours de la description des applications pour la veille, les services Web 2.0 ne sont pas suffisants en soit pour la recherche d'information. De plus, bien que ses possibilités soient très puissantes, l'utilisation du Web 2.0 reste encore marginale à l'échelle de tous les internautes : Une étude du cabinet *Pew Internet*¹⁵⁷ montre que l'activité internet dominante, sur une journée type d'un américain reste tout simplement le mail : 53% des internautes l'utilisent. Ils sont 38% à utiliser un moteur de recherche. Par contre, ils ne sont que 8% à créer des blogs et 26% à partager des photos et vidéos, cette collaboration multimédia étant de loin l'activité la plus courante du Web 2.

Une enquête en ligne de l'Université d'Oxford réalisée auprès de 1469 personnes (une majorité d'étudiants) fin 2006 détaille les usages du Web 2.0 ¹⁵⁸. Elle révèle elle aussi que les applications de ce web nouvelle génération sont encore peu connues et peu utilisées : La majorité des répondants n'a jamais employé ou jamais entendu parlé des fils RSS ou des applications de *Social Bookmarking* comme *Del.icio.us*.

Aussi, l'influence grandissante du Web 2.0 ne doit pas masquer les usages "traditionnels" du web qui restent majoritaires.

Une population souvent isolée et non représentative

L'âge

Si de nombreuses applications "sociales" apportent un réel progrès, il faut tempérer la nature de certains contenus issus du mode collaboratif: celui-ci n'est en effet pas représentatif des internautes puisque le public du Web 2.0 est une population plutôt jeune.

Ainsi, l'étude *Pew Internet* révèle que l'audience de Wikipedia aux USA était en août 2006 à 47% située dans la tranche des 18-34 ans. Les plus de 45 représentaient 28%. Sur le site d'*Encarta*, l'encyclopédie de Microsoft plus traditionnelle (Web 1.0), la part des plus de 45 ans atteint 33%.

Des communautés isolées, structurées en îles

Les communautés actuelles du Web 2.0 sont souvent renfermées sur elles-mêmes : les statistiques sur les liens pointant vers les blogs et sur les liens présents dans les billets révèlent que les blogueurs se lient et se lisent beaucoup entre eux. L'internaute Web 2.0 aurait donc tendance à s'enfermer sur lui-même et autour d'individus fédérés en communautés focalisés sur les mêmes centres d'intérêt que lui.

Par ailleurs, l'hyper segmentation et le ciblage des nouvelles rendu possible par le flux RSS notamment, peut conduire à favoriser ces microcosmes communautaires : Parcourir un journal papier peut favoriser davantage l'ouverture, la prise de recul et la mise en perspective des différentes informations que la lecture de flux aux thématiques très spécialisées.

Les différentes communautés, par réflexe et de par la nature des applications Web 2.0 qui se dupliquent mais ne se connectent pas, ont tendance à peu communiquer entre elles. Ainsi, les liens entre réseaux sociaux sont rares. Peu de personnes

enregistrés sur LinkedIn iront également s'inscrire sur Viadeo ou FaceBook. L'atomisation des données sur des dizaines d'applications reste également un vrai problème.

Des domaines mal couverts

Certaines thématiques vont être surreprésentées au détriment d'autres au traitement beaucoup plus confidentiel car boudées par les internautes producteurs volontaires d'information. Les ingénieurs, techniciens et jeunes étant fortement représentés parmi les blogueurs, les thématiques high tech sont très largement couvertes par les blogs : par exemple, dans le domaine de la RFID (Identification par Radio Fréquence), on compte plus de 100 blogs, idem pour la VoIP (Voix sur IP). A l'inverse, si vous travaillez dans le secteur des machines outils ou de la plasturgie, il vous sera difficile de trouver du contenu sur la blogosphère.

Tous les secteurs ne sont donc pas égaux en volume sur le net et davantage encore sur la blogosphère et sur les autres applications collaboratives où tout internaute peut publier son propre contenu et où certaines tranches d'âges et catégories professionnelles ne sont pas bien représentées.

Conclusion

Le Web 2.0 : ce mot qui génère énormément de buzz, d'écrits et de blogs spécialisés n'est-il qu'un terme marketing, un courant de mode ¹⁵⁹ ? Il est probable que bientôt, on ne l'emploiera plus pour désigner ces nouvelles applications qui seront devenues la norme. Toutefois, ce terme, sorte de parapluie conceptuel qui recouvre les différentes tendances du web nouvelle génération, a le mérite de mettre un nom simple sur un courant du net novateur.

Le Web 2.0 n'est pas une révolution ou une rupture technologique mais plus certainement une évolution qui permet de replacer l'utilisateur au centre des réseaux. Cela n'est certes pas très nouveau : Usenet, l'un des plus anciens protocoles de l'internet, bien antérieur au WWW, offrait du contenu en ligne créé par l'internaute via les Newsgroups dès les années 1980.

Mais les progrès technologiques ont permis d'intensifier cette tendance et de démocratiser l'accès à la création de contenu en ligne et à sa diffusion au plus grand nombre. Potentiellement, tout internaute peut maintenant diffuser ses publications tant les outils sont devenus simples.

Le Web 2.0 tend à s'insérer au cœur du quotidien de chacun en lui apportant des services novateurs réellement utiles et qui rendent possible l'apparition de nouveaux usages. Ainsi, les mash-up qui s'appuient sur la géolocalisation ouvrent des possibilités inenvisageables jusqu'ici. On peut même imaginer à terme une superposition de ces cartes dynamiques à la vision réelle d'un piéton ou d'un automobiliste. Le Web 2.0 est en perpétuelle évolution, ses capacités d'innovation sont très fortes car elles reposent souvent sur des recombinaisons audacieuses de l'existant, hybridant applications et services dans un système de développement "open source" ... L'innovation est aussi innovation de postures (prendre la parole, consommer du contenu "amateur"...), qui parfois challengent l'ordre établi et redistribuent les "pouvoirs". La position des acteurs du Web 1.0 ou des médias classiques pourrait se trouver de plus en plus remise en question là où les internautes tendent à se placer en experts.

La place de l'humain

Si celle-ci est devenue prépondérante, il est toutefois difficile de s'affranchir d'organisations un tant soit peu hiérarchisées, l'autorégulation, voir Wikipedia, ayant ses limites.

En outre, la question de la consommation de contenu et de sa qualification sur l'internet devient cruciale. La pertinence des votes de tous les internautes est discutable, ou du moins, le fait qu'ils puissent à eux seuls qualifier et évaluer la qualité. Par ailleurs, que représente exactement un vote : une adhésion au contenu, à la forme, aux idées, aux écrits d'un ami, à une idée avec laquelle je suis en désaccord mais qui à le mérite d'exister ?

Finalement, il faut continuer à normer, organiser, coordonner, réguler...notamment pour éviter les diffamations ou la désinformation. Pour ce qui est du contenu et de sa qualité, les critiques émises à l'encontre de Wikipédia nous incitent à étudier de plus près les initiatives qui hybrident les interventions d'amateurs avec les compétences de professionnels.

Nous sommes passés en 10 ans d'un contenu généré par quelques internautes pour la masse à un contenu de masse à destination de la masse ("Everybody to Everybody"). Cela ne peut guère se développer sans des mécanismes qui replacent les internautes les plus spécialisés dans des processus de validation d'une information devenue pléthorique.

Les données personnelles

Simplicité, réactivité, personnalisation, souplesse, mutualisation...font partie des avantages les plus appréciables du web 2.0. Mais cette liberté de publication et de diffusion n'est elle pas parfois illusoire ? Les utilisateurs doivent rester vigilants face aux problèmes de confidentialité, sécurité et traçabilité de l'information¹⁶⁰. Des voix s'élèvent pour dénoncer les risques liés à la maîtrise des données personnelles, ou au respect des droits d'auteurs. A noter que les licences *Creative Commons*¹⁶¹ permettent de mieux définir les contours de cette liberté, à partir de contrats flexibles de droit d'auteur. Ce système *Creative Commons* donne la possibilité, gratuitement, et simplement aux titulaires de droits de choisir et d'exprimer les conditions d'utilisation de leurs œuvres, à travers onze possibilités. Quant aux utilisateurs, ils n'ont pas à négocier systématiquement une autorisation avant toute utilisation ou réutilisation.

Un autre problème réside aussi dans l'éparpillement des données dans des sites différents ou parfois redondants. Il s'agira donc de définir les objets (livre, personne, CD, photos...) à partir d'un identifiant. *"Cette bataille pour savoir qui gèrera ces données de base telles que la position géographique, les espaces de noms et le code unique de chaque produit a déjà commencé"* estime Tim O'Reilly ¹⁶².

Les perspectives d'évolution vers un "Web 3"

Le web actuel est déstructuré et la classification et l'organisation de son contenu posent problème. La recherche de nouveaux modes de recherche et de structuration est une sorte de quête du graal que de nombreux experts et entreprises poursuivent.

La folksonomie repose sur une idée originale qui sous tend toute l'idéologie Web 2.0 : celle de l'intelligence collective et de la décentralisation. Mais les limites de la folksonomie appelée aussi "web sémantique du pauvre" remettent en question l'affirmation selon laquelle l'intelligence collective et la libéralisation totale du fonctionnement du web 2.0 sont une solution viable.

Aussi, la nécessité d'un Web sémantique ¹⁶³ fiable se fait de plus en plus urgente. Cette nouvelle phase de développement du Web permettrait d'améliorer l'exploitation des informations par les utilisateurs en obtenant une information organisée, normalisée de façon à rendre le web compréhensible aux machines pour faciliter l'exploitation par les utilisateurs.

Dans un article paru dans le New York Times en novembre 2006, John Markoff imagine un web 3.0 selon le principe *"d'ajout d'une couche de sens tout en haut du web existant, pour en faire un guide au lieu d'un simple catalogue"*. Pour l'auteur, *"nous sommes en train de passer d'un web de documents connectés à un web de données connectées"*.

Le web sémantique offrira un langage commun permettant aux données de renvoyer à des objets du monde réel d'une façon identique. Le projet *Freebase*, lancé par la start up *Metaweb*, illustre cette approche : il a pour objectif de faciliter la recherche et le partage d'informations à travers une base de connaissances communes, structurées via des métadonnées permettant de décrire chaque élément d'information. Ces éléments seraient reliés entre eux par un système de relations. Tim O'Reilly définit le projet comme *"la machine à créer les synapses du cerveau global"*.

Il s'agirait de *"créer en quelque sorte un Google du web sémantique, un répertoire global d'informations structurées, intelligibles et exploitables autant par des machines que par*

des humains -et donc capables d'extraire et d'interpréter les relations entre des données pour, par exemple, répondre à des requêtes par des solutions plutôt que des listes de document " ¹⁶⁴.

La veille

Dans cet univers, le chercheur d'informations et le veilleur doivent...garder la tête froide : la forte médiatisation du Web 2.0 ne doit pas masquer la pertinence d'outils plus traditionnels. Les différentes applications Web 2.0, nous l'avons vu, suffisent rarement pour mener à bien une collecte d'information. Certes, les bookmarks sociaux apportent un contenu sélectionné par l'internaute mais on ne peut se départir des moteurs de recherche web classiques. La surveillance des flux RSS permet d'être alerté sur de nombreux sujets mais la majorité des sites web restent encore dépourvus de ces fils. Les tags apportent une véritable souplesse à l'utilisateur mais la taxonomie des bibliothèques ou des répertoires web- pour ne citer que ces lieux- reste essentielle pour se repérer à travers des milliers de ressources.

Le Web 2.0 pour la recherche et la veille n'en est qu'à ses balbutiements. Aussi, avant qu'il ne "digère" complètement, par enrichissement, les ressources et les applications Web 1.0, il convient de faire coexister ces "2 web" dans les pratiques.

Le Web 2.0 apporte une richesse essentielle : l'expression d'individus porteurs d'idées et de facultés susceptibles d'apporter des changements, de re-questionner l'"ordre établi", éléments indispensables à l'évolution de la culture à condition de ne pas oublier que des volontés ambitieuses doivent intervenir pour canaliser les initiatives sans les étouffer et leur insuffler une véritable ambition. Voilà tout l'enjeu...

A propos des auteurs

DIGIMIND SERVICES

Digimind Services accompagne les plus grandes entreprises pour la conception, la mise en œuvre et le déploiement de projets de veille qui reposent sur la solution Digimind Evolution, apportant à ses clients un retour sur investissement de plus de 600%, et ce, dès la première année. Les méthodologies propriétaires du département conseil ainsi que son expertise des problématiques et sources d'informations sur plus de 20 secteurs d'activité, développée auprès de ses clients depuis de nombreuses années, permet aux entreprises d'anticiper les changements de leur environnement pour prendre les meilleures décisions sur leur marché.

Le conseil Digimind porte sur tous les aspects du workflow de veille : ciblage stratégique, collecte des informations, traitement et analyse, exploitation et diffusion, gestion de projet, conseil organisationnel et gestion du changement, formations et support fonctionnel et technique.

CHRISTOPHE ASSELIN

Issu d'un cabinet d'études de marché B2B et spécialisé depuis 1997 dans la mise en place de systèmes de veille (e-France.org, Ecole Militaire), Christophe Asselin allie une connaissance approfondie des secteurs économiques et une parfaite maîtrise des outils de recherche sur internet et des solutions avancées de veille. Expert reconnu, il édite le site <http://www.intelligence-center.com> et le blog <http://influx.joueb.com>, tout deux consacrés à la recherche d'information.

Spécialiste de la veille internet chez Digimind, il accompagne les clients dans la mise en place de leur dispositif de veille (expression des besoins, définition de plans de veille, sourcing, architecture, paramétrage, formation, accompagnement). Il intervient ainsi auprès de sociétés dans différents secteurs : les télécoms avec France Telecom R&D, l'industrie pharmaceutique avec Sanofi Aventis, Roche Pharma, Expanscience ainsi qu'Alstom Transport, Veolia Environnement et des références confidentielles dans le secteur des biotechnologies, de l'énergie, des télécoms, de l'imprimerie, de la défense, du conseil...

[Merci au Service Marketing, à Malaïac www.malaiac.net et à Elsa Bonier, étudiante en Master 2 Management des innovations en communication à Paris XIII, pour leur fructueuse collaboration].

VERONIQUE MESGUICH

Après avoir fondé et co-dirigé pendant six ans un cabinet spécialisé en veille technologique, Véronique Mesguich dirige l'Infothèque du Pôle Universitaire Léonard de Vinci. Elle enseigne également la maîtrise de l'information stratégique dans plusieurs établissements d'enseignement supérieur, et anime régulièrement des sessions de formation continue autour de la méthodologie de recherche avancée et de veille sur Internet.

Véronique Mesguich est l'auteur, avec Armelle Thomas, de l'ouvrage "Net recherche : le guide pour trouver l'information utile", dont la deuxième édition est parue en mars 2007 aux éditions de l'ADBS.

Webographie Bibliographie

Sources d'information sur le Web 2.0

Sites web, wikis et blogs

• Répertoires et classement de services et sociétés Web 2.0

- 2.0 Websites

<http://2.0websites.com/>

- Web 2.0 Magazine : Top 100 Web 2.0 Sites

<http://web2magazine.blogspot.com/2007/01/thanks-for-web-2.html>

- Web 2.0 Magazine : The Best Web 2.0 Software of 2006

http://web2.socialcomputingmagazine.com/the_best_web_20_software_of_2006.htm

- Complete List of Web 2.0 Products and Services

<http://www.listible.com/list/complete-list-of-web-2-0-products-and-services>

- The Complete Web 2.0 Directory

<http://www.go2web20.net/>

- FredCavazza.net : le meilleur du Web 2.0

<http://www.fredcavazza.net/index.php?2005/12/12/986-le-meilleur-du-web-20>

- Web 2.0 Directory : 1200+ Web 2.0 Sites in 50+ categories

<http://web2.econsultant.com/index.html>

- Top 25 Web 2.0 Search Engines

<http://oedb.org/library/features/top-25-web20-search-engines>

- 25 start-ups Web 2.0 prometteuses

http://money.cnn.com/2007/02/21/magazines/business2/nextnet_intro.biz2/

▪ **Blogs et sites spécialisés**

- Le groupe de travail Web2.0 francophone
<http://www.web20workgroup.fr/>
- Web 2.0 Work Group
<http://web20workgroup.com>
- Dessine-moi le Web 2.0
<http://www.deuxzero.com/>
- FredCavazza.net
<http://www.fredcavazza.net/>
- TechCrunch en français
<http://fr.techcrunch.com/>
- TechCrunch
<http://www.techcrunch.com>
- Le web 2.0, c'est pas du buzz
<http://web2rules.blogspot.com/>
- Mashable France
<http://fr.mashable.com/>
- Mashable
<http://mashable.com/>
- AccessoWeb
<http://blog.accessoweb.info/>
- Web 2.0 Magazine :
<http://web2magazine.blogspot.com>
- Révolution Web 2.0 en Live !
<http://www.webdeux.info/>
- Read/WriteWeb
<http://www.readwriteweb.com/>
- Web 2.0 Innovations
<http://web2innovations.com/>
- Web 2.0 Summit
<http://www.web2con.com>
- Original Signal (agrégateur de flux RSS)
<http://www.originalsignal.fr/>
- What's Web 2.0 (agrégateur de flux RSS)
<http://www.whatsweb20.com/aggregator?from=100>
- Le blog de Bloxx : Communiquons avec le Web 2.0
<http://blog.bloxx.fr/>

- FrenchWeb.org : Le blog des professionnels français du Web 2.0
<http://www.frenchweb.org/>

- Blog Web 2.0 et Marketing
<http://www.vincentabry.com/>

- Le Web 2.0 en Action
<http://leweb2.blogspot.com/>

- aYsoon 2.0
<http://blog.aysoon.com/>

Emmanuel de Taillac
<http://emmanueldetaillac.fr/>

- NiouMedia
<http://nioumedia.com/>

- 2803
<http://www.2803.com/>

- Web 2 francophone (Wiki)
<http://web2fr.com/Accueil>

▪ **Blogs et sites avec un axe Web 2.0**

- Publishing 2.0
<http://publishing2.com/>

- Social-Computing
<http://www.social-computing.com/>

- Ecosphere
<http://ecosphere.wordpress.com/>

- inFLUX
<http://influx.joueb.com>

- Benoit Descary
<http://descary.com/>

- Jean Michel Billaut
<http://billaut.typepad.com/jm/>

- Blog francophone des TIC
<http://www.2607.fr/>

20Minutes : Top 50 tech et médias
<http://www.20minutes.fr/article/150876/20070410-Blogs-Top-50-tech-et-media.php>

- Intruders.tv
http://new.intruders.tv/intruders_fr/

- Jeff Clavier "Sans Accent"
<http://sansaccent.softtechvc.com/>
- Demain tous journalistes ?
<http://benoit-raphael.blogspot.com>
- Collectif l'ouvre-boîte : Initiatives sur le partage d'information
<http://louvre-boite.viabloga.com/>
- Zero seconde
<http://zeroseconde.blogspot.com/>

Articles

▪ Introduction au Web 2.0

- Web 2.0: l'article de O'Reilly résumé pour les fainéants...
<http://www.expressions.be/2005/11/01/web-20-resume-pour-les-faineants/>
- Dossiers Web 2.0 - Journal Du net
<http://www.journaldunet.com/dossiers/web2.0/>
- Avis des managers : le Web 2.0 à la loupe
<http://www.journaldunet.com/solutions/0603/diaporama/avis-managers-web2/1.shtml>
- Les entreprises face aux défis du Web 2.0
<http://www.journaldunet.com/solutions/0704/070413-entreprises-web-20/1.shtml>
- The Machine is Us/ing Us Transcription (vidéo)
<http://mediatedcultures.net/ksudigg/?p=78>
- The New Wisdom of the Web
<http://www.msnbc.msn.com/id/12015774/site/newsweek/>
- Qu'est-ce que Web 2.0 ?
<http://www.01net.com/article/294819.html>
- Approaching a definition of Web 2.0 :
<http://socialsoftware.weblogsinc.com/2005/09/29/approaching-a-definition-of-web-2-0/>
- Fonctionnement de l'entreprise classique vs entreprise 2.0
<http://alban.cornillet.free.fr/akw/index.php/2006/10/02/fonctionnement-de-lentreprise-classique-vs-entreprise-20/>
- Médias traditionnels et acteurs du Web 2.0 : vers la cohabitation ou la convergence des acteurs de l'information et du divertissement ?"
<http://www.scribd.com/doc/50818/These-sur-le-Web-2-Cabanis>

▪ **Ajax**

- L'Ajax en 5 questions

http://www.journaldunet.com/solutions/0602/060208_qr-ajax.shtml

- Ajax et Moteurs 2.0

<http://www.moteurzine.com/archives/2006/moteurzine127.html#2>

▪ **Les Mashups**

- Les Mashups, applications Web composites très Web 2.0

<http://www.journaldunet.com/solutions/0606/060628-qr-mashup-hybride.shtml>

- Le modèle économique des Mashups

<http://www.internetactu.net/?p=6285>

▪ **Bookmarks sociaux, Social search**

- Le Social Bookmarking au secours de la pertinence de la recherche d'information

<http://www.gillesbalmisse.com/v2/spip.php?article171>

- What's the Big Deal With Social Search?

<http://searchenginewatch.com/showPage.html?page=3623153>

- Who's Who in Social Search

<http://searchenginewatch.com/showPage.html?page=3623173>

Ouvrages

- Web 2.0 - Les internautes au pouvoir - Jean-François Gervais

<http://www.dunod.com/pages/ouvrages/ficheouvrage.asp?id=50701>

- Les réseaux sociaux, Pivot de l'internet 2.0 - Alain Lefebvre,

http://www.cluster21.com/fr/bookreview/livre_les_reseaux_sociaux

- Web 2.0 : (R)évolutions et nouveaux services d'Internet - Jean-Noël Anderruthy

http://www.amazon.fr/Web-2-0-%C3%A9volutions-nouveaux-dInternet/dp/2746034972/ref=sr_l_18/402-7140143-8118507?ie=UTF8&s=books&qid=1178118687&sr=8-18

- Web 2.0 : The Future of the Internet and Technology Economy and How Entrepreneurs, Investors, Executives & Consumers Can Take Ad

http://www.amazon.com/Web-2-0-Technology-Entrepreneurs-Executives/dp/1587622009/ref=sr_l_6/103-0331542-4028646?ie=UTF8&s=books&qid=1178118732&sr=1-6

- La Longue Traîne : La nouvelle économie est là ! - Chris Anderson

<http://www.amazon.fr/Longue-Tra%C3%A9ne-nouvelle-%C3%A9conomie->

[est/dp/2744062693/ref=sr_l_1/403-1740173-8750801?ie=UTF8&s=books&qid=1180542516&sr=1-1](http://www.amazon.com/Geospatial-Web-Geobrowsers-Information-Processing/dp/1846288266/ref=sr_l_1/403-1740173-8750801?ie=UTF8&s=books&qid=1180542516&sr=1-1)

- The Geospatial Web: How Geobrowsers, Social Software and the Web 2.0 are Shaping the Network Society

http://www.amazon.com/Geospatial-Web-Geobrowsers-Information-Processing/dp/1846288266/ref=sr_l_28/103-0331542-4028646?ie=UTF8&s=books&qid=1178118760&sr=1-28

- Les Wikis, espaces de l'intelligence collective - Jérôme Delacroix

http://www.amazon.fr/Wikis-J%C3%A9r%C3%B4me-Delacroix/dp/2952051445/ref=sr_l_1/403-1740173-8750801?ie=UTF8&s=books&qid=1180544723&sr=8-1

- Pro Web 2.0 Mashups: Remixing Data and Web Services

http://www.amazon.fr/Pro-Web-2-0-Mashups-Remixing/dp/159059858X/ref=sr_l_4/403-1740173-8750801?ie=UTF8&s=english-books&qid=1180542751&sr=8-4

- Web 2.0 et Web avancé - Alexandre Boni, Nicolas Stemart

<http://www.amazon.fr/Web-2-0-avanc%C3%A9-Alexandre-Boni/dp/2742967281>

Dans la même collection

- White Paper
"Blogs et RSS, des outils pour la veille stratégique"
Christophe Asselin , Consultant Veille Stratégique, Digimind

 - White Paper
"Découvrir et exploiter le Web Invisible pour la veille stratégique"
Christophe Asselin , Consultant Veille Stratégique, Digimind

 - White Paper
"Catégorisation automatique de textes"

 - White Paper
"Benchmark des solutions de veille stratégique"

 - White Paper
"Evaluer le Retour sur Investissement d'un logiciel de veille"
Edouard Fillias, Consultant Veille Stratégique, Digimind
- > **A télécharger sur** http://www.digimind.fr/services/white_paper.htm

Notes

-
- ¹ http://www.guim.fr/blog/2006/12/le_film_autumn_.html
<http://www.lemonde.fr/web/article/0,1-0@2-651865,36-875106,0.html>
- ² <http://video.google.fr/videoplay?docid=724218352255575783>
- ³ Map of online communities par Randall Munroe <http://www.xkcd.com/c256.html>
- ⁴ <http://www.deuxzero.com/>
- ⁵ <http://www.paulgraham.com/bio.html>
http://fr.wikipedia.org/wiki/Paul_Graham
- ⁶ Qu'est ce que le web 2.0. Hubert Guillaud, Internet Actu, 29 mai 2005 <http://www.internetactu.net/?p=6144>
- ⁷ <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
<http://web2rules.blogspot.com/2006/01/what-is-web-20-par-tim-oreilly-version.html>
- ⁸ <http://www.zdnet.fr/actualites/internet/0,39020774,39334479,00.htm>
- ⁹ <http://technology.guardian.co.uk/weekly/story/0,,1823959,00.html?gusrc=rss>
- ¹⁰ <http://www.internetactu.net/?p=7035>
- ¹¹ <http://www.hitwise.com/datacenter/>
- ¹² http://ross.typepad.com/blog/2006/04/power_law_of_pa.html
- ¹³ http://blogs.forrester.com/charleneli/2007/04/forresters_new_.html
- ¹⁴ <http://www.tagcrowd.com>
- ¹⁵ "You Say You Want a Web Revolution" <http://www.wired.com/science/discoveries/news/2005/08/68403>
- ¹⁶ <https://addons.mozilla.org/fr/firefox/addon/748>
<http://www.journaldunet.com/developpeur/tutoriel/out/050525-greasemonkey-presentation.shtml>
- ¹⁷ Mashups Google Maps :
<http://www.mibazaar.com/extras/fortune100.html>
<http://www.weatherbonk.com/weather/index.jsp> http://gmaps.tommangan.us/spacecraft_tracking.html
- ¹⁸ <http://www.google.com/calendar>
- ¹⁹ <http://docs.google.com/>
- ²⁰ <http://www.moovement.fr/>
- ²¹ <http://www.cooptin.com/fr/WEB/home.aspx>
<http://www.jobmeeters.com/fr/accueil/index.php>
- ²² <http://www.zlio.com/>
<http://shopping.yahoo.com/shopsphere/>

²³ <http://www.ciao.fr/>
<http://www.wantuno.com/>

²⁴ <http://maps.google.fr/>
<http://www.geoportail.fr>
<http://earth.google.fr/>
<http://maps.yahoo.com/>
<http://www.flickr.com>
<http://www.zdnet.fr/actualites/internet/0,39020774,39366354,00.htm>

²⁵ <http://www.criteo.com/>

²⁶ From Youtube to Youniversities, Henry Jenkins, 16 février 2007
http://www.henryjenkins.org/2007/02/from_youtube_to_youniversity.html

²⁷ <http://affordance.typepad.com/>

²⁸ <http://www.plosone.org/home.action>

²⁹ Web collaboratif, web coopératif, web 2.0, quelles interrogations pour l'enseignant documentaliste, Marie France Blanquet, 12 janvier 2007 <http://eprofsdocs.crdp-aix-marseille.fr/IMG/rtf/intervention-MF-Blanquet.rtf>

³⁰ <http://veillepedagogique.blog.lemonde.fr/>

³¹ <http://bibliobsession.free.fr/dotclear/>

³² http://liblogs.albany.edu/library20/2006/11/a_librarians_20_manifesto.html

³³ <http://del.icio.us/>

³⁴ <http://www.connotea.org/>
<http://www.citeulike.org/>
<http://www.snipitron.com>

³⁵ <http://del.icio.us/>
<http://www.blinklist.com/>
<http://blogmarks.net/>
<http://www.furl.net/>
<http://fr.beta.bookmarks.yahoo.com>
<http://myweb2.search.yahoo.com/>

³⁶ <http://www.stumbleupon.com/>
<http://similicio.us/>
<http://www.yoono.com/index.jsp>

³⁷ Social bookmarks' charts version 3.5
http://www.irox.de/file_download/3
<http://www.roxomatic.de/856>

Social Bookmarking Services And Tools: The Wisdom Of Crowds That Organizes The Web
http://www.masternewmedia.org/news/2006/12/01/social_bookmarking_services_and_tools.htm

Free Bookmark Managers

<http://www.feedbus.com/bookmarks/>

³⁸ <http://www.flickr.com/>

³⁹ <http://www.flickr.com/groups/tehi/pool>

⁴⁰ <http://www.flickr.com/photos/tags/northamericaninternationalautoshow>

⁴¹ <http://photobucket.com>

⁴² <http://www.squidoo.com>

⁴³ <http://www.squidoo.com/nanotechRev>

⁴⁴ <http://www.squidoo.com/tags/marketing>

⁴⁵ <http://www.youtube.com/>

⁴⁶ <http://www.dailymotion.com>

⁴⁷ Une sélection de moteurs de recherche vidéo et de services de partage et téléchargement de vidéos.

<http://c.asselin.free.fr/french/videos.htm>

⁴⁸ <http://www.emarketer.com/Article.aspx?id=1004918>

⁴⁹ <http://www.viadeo.com>

⁵⁰ <http://www.linkedin.com/>

⁵¹ <http://www.6nergies.net/>

<http://www.xing.com/>

⁵² <http://www.facebook.com/>

⁵³ http://en.wikipedia.org/wiki/List_of_social_networking_sites

⁵⁴ <http://www.linkrank.com/business.php>

⁵⁵ <http://digg.com>

⁵⁶ <http://www.nuouz.com>

<http://www.blogmemes.com>

<http://www.scoopeo.com/>

<http://www.fuzz.fr/>

<http://www.tapemoi.com>

<http://meneame.net/>

<http://www.rojo.com/>

<http://www.beta.netscape.com/>

⁵⁷ <http://www.wikio.fr/>

⁵⁸ <http://www.agoravox.fr>

⁵⁹ Carlo Revelli, Joël de Rosnay

http://fr.wikipedia.org/wiki/Carlo_Revelli

http://fr.wikipedia.org/wiki/Jo%C3%ABl_de_Rosnay

⁶⁰ Agoravox : politique éditoriale http://www.agoravox.fr/article.php3?id_article=60

⁶¹ <http://www.readspeaker.com/>

⁶² <http://english.ohmynews.com>

⁶³ State of the Blogosphere- avril 2007. <http://www.sifry.com/alerts/archives/000493.html>

⁶⁴ Les blogs dans le monde. Estimation à partir des chiffres produits par :
Digimind, Blog Herald, Technorati , Pew/Internet, Médiamétrie, Loïc le Meur, Live Journal, Skyblogs, Overblogs, Terra, MSN Spaces, Splinder, Excite: Bloggers.it, Timblog, W3B, Skynet, Blog.hr, Overskrift.dk, Smartlog, onet.pl, Tenbit mylog.pl, eblog.pl, Blog.pl, Blogi.pl, blog.central.is/, Romandie.fr, Blogspot, timsoft.ro, Globo.com, Financial Express, Prixadi blog, Baidu, Yahoo Blogs South Korea, Planet Weblog Service, Cyworld, L'Atelier et ZDNet Australia

⁶⁵ <http://www.nature.com/webfeeds/index.html> <http://www.sciencemag.org/rss/>

⁶⁶ <http://search.live.com>

⁶⁷ <http://search.yahoo.com/>

⁶⁸ <http://news.google.fr/>
<http://fr.news.yahoo.co>
<http://fr.newsbot.msn.com/>
<http://www.topix.net>
<http://www.wikio.fr/>

⁶⁹ <http://groups.google.fr/group/rec.travel.europe/feeds?hl=fr>

⁷⁰ <http://www.forummobiles.com/index.php?showforum=200>

⁷¹ <http://fr.groups.yahoo.com/>

⁷² <http://fr.myweb2.search.yahoo.com>
<http://blogmarks.net/>
<http://del.icio.us/>

⁷³ <http://www.flickr.com/>
<http://photobucket.com>

⁷⁴ <http://www.youtube.com/>
<http://www.dailymotion.com/fr>

⁷⁵ "Different Engines, Different Results, Web Searchers Not Always Finding What They're Looking for Online".
<http://comparesearchengines.dogpile.com/OverlapAnalysis.pdf>

⁷⁶ <http://blogsearch.google.fr/>
<http://fr.ask.com/#subject:bls|pg:1>
<http://www.feedster.com/>
<http://www.blogpulse.com/>
<http://www.technorati.com/>
<http://www.sphere.com/>

⁷⁷ <http://www.sifry.com/alerts/archives/000493.html>

⁷⁸ <http://fr.search.yahoo.com/web/advanced?ei=UTF-8&ybs=0>

<http://www.exalead.fr/search/results?q=sarkozy&%24mode=allweb>

⁷⁹ <http://www.blogarama.com/>
<http://www.rss-network.com/>

<http://www.rssfeeds.com/>

⁸⁰ <http://www.43things.com/>
<http://36trucs.com/>

⁸¹ http://trainque.com/code/delicious_lookup.php

⁸² <http://technorati.com/tag>

⁸³ <http://www.searchles.com>

⁸⁴ <http://tagcentral.net>
<http://www.keotag.com/>
<http://www.thagoo.com>
<http://www.zewol.net/>
<http://www.gnosh.org>

⁸⁵ <http://www.zibb.com/>

⁸⁶ <http://www.ecrans.fr>

⁸⁷ Google Tag Cloud Maker <http://userscripts.org/scripts/show/6171>

⁸⁸ <http://newzingo.com/>

⁸⁹ Folksonomies, les usagers indexent le web, Olivier le Deuff, Bulletin des Bibliothèques de France, n°4, 2006

⁹⁰ Folksonomy, les tags en délire, Cyril Fiévet, Internetactu, 13 octobre 2005
<http://www.internetactu.net/?p=6166>

⁹¹ <http://www.zoominfo.com>

⁹² <http://www.ziki.com/>

⁹³ <http://www.yoname.com>
<http://www.stalkerati.com/>
<http://wink.com>

⁹⁴ <http://fr.answers.yahoo.com/>
<http://answers.yahoo.com/>
<http://ask.yahoo.com/>

⁹⁵ <http://answers.google.com/answers/>

⁹⁶ <http://iq.lycos.fr/>
<http://qna.live.com/beta.aspx>

⁹⁷ <http://www.experts.com/directory.asp>

⁹⁸ http://search.wikia.com/wiki/Search_Wikia

⁹⁹ <http://decipho.com/>

¹⁰⁰ <http://www.yoono.com>

¹⁰¹ <http://www.foox.com>

<http://www.webzlle.com>

¹⁰² <http://influx.joueb.com/news/powerset-et-xerox-parc-s-attaquent-a-google-naturellement>

¹⁰³ <http://www.medgle.com/>
<http://www.medworm.com>
<http://www.medstory.com/>
<http://www.searchmedica.co.uk/>

¹⁰⁴ <http://www.rollyo.com/>
<http://swickihome.eurekster.com/>
<http://google.com/coop/cse/overview>
<http://builder.search.yahoo.com/>

¹⁰⁵ "Different Engines, Different Results, Web Searchers Not Always Finding What They're Looking for Online".
<http://comparesearchengines.dogpile.com/OverlapAnalysis.pdf>

¹⁰⁶ Le "web invisible" (deep web, hidden web) désigne la partie du web non accessible aux moteurs de recherche classiques. Le web invisible comprend des bases, banques de données et bibliothèques en ligne gratuites ou payantes...Il est constitué de pages web "mal" ou non indexées par les moteurs.

Pour plus de précisions voir le White Paper : "Découvrir et exploiter le Web Invisible pour la veille stratégique"
http://www.digimind.fr/services/white_paper.htm

¹⁰⁷ <http://www.thelancet.com>

¹⁰⁸ <http://www.digimind.fr/produits/finder.htm>

¹⁰⁹ <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=pubmed>

¹¹⁰ <http://pmi.nlm.nih.gov/interact/index.php#top> <http://www.hubmed.org>

¹¹¹ <http://www.uspto.gov/patft/index.html>
<http://www.washingtonpost.com/wp-dyn/content/article/2007/03/04/AR2007030401263.html>

¹¹² <http://www.google.com/history>

¹¹³ <http://www.google.com/history/trends>
<http://www.google.com/history/items>

¹¹⁴ http://fr.wikipedia.org/wiki/Application_composite

¹¹⁵ <http://www.mibazaar.com/extras/fortune100.html>

¹¹⁶ <http://tagmaps.research.yahoo.com/worldexplorer.php>

¹¹⁷ <http://flickrvision.com/>

¹¹⁸ http://gmaps.tommangan.us/spacecraft_tracking.html

¹¹⁹ <http://www.healthmap.org/>
<http://www.brevardbiodiesel.org/map/bd.html>
<http://kartametro.info/karta/default.aspx?hl=EN>
<http://www.wikimapia.org/>
<http://www.globalincidentmap.com/home.php>
<http://dartmaps.mackers.com/>

-
- ¹²⁰ <http://www.tagurls.com/>
- ¹²¹ <http://www.themolu.com/>
<http://www.gnosh.org/>
- ¹²² <http://www.webmashup.com/>
<http://www.programmableweb.com/mashups/directory>
<http://mashupfeed.com/>
- ¹²³ <http://googlemapsmania.blogspot.com>
- ¹²⁴ <http://local.alkemis.com>
- ¹²⁵ <http://www.vox.com/>
<http://www.over-blog.com/>
<http://www.blogspirit.com/fr/index.php>
<http://www.typepad.com/>
<https://www.blogger.com/start>
<http://viabloga.com/>
<http://www.hautetfort.com/>
- ¹²⁶ <http://c.asselin.free.fr/french/weblogs.htm#outils>
- ¹²⁷ <http://www.clubic.com/article-67515-8-comparatif-plateformes-blogs.html>
http://www.ojr.org/ojr/images/blog_software_comparison.cfm
- ¹²⁸ <http://www.feedburner.com/fb/a/home>
- ¹²⁹ <http://pages.google.com/>
- ¹³⁰ http://www.digimind.fr/services/white_paper.htm
- ¹³¹ <http://fr.wikipedia.org/wiki/Accueil>
- ¹³² <http://blog.outter-court.com/archive/2006-12-05-n22.html>
- ¹³³ <http://www.nature.com/news/2005/051212/full/438900a.html>
<http://www.nature.com/nature/britannica/index.html>
http://www.nature.com/news/2005/051212/multimedia/438900a_m1.html
- ¹³⁴ <http://www.wikiseek.com/>
- ¹³⁵ <http://www.exalead.fr/>
- ¹³⁶ <http://www.wikimindmap.org>
- ¹³⁷ http://en.citizendium.org/wiki/Main_Page
- ¹³⁸ <http://www.alacrawiki.com>
- ¹³⁹ <http://ceclines.xwiki.com/xwiki/bin/view/Main/Fils+RSS>
- ¹⁴⁰ <http://www.wikidweb.com/>
- ¹⁴¹ <http://www.opera.com/> <http://www.maxthon.com/>
<http://www.microsoft.com/france/windows/ie/>
<http://www.mozilla-europe.org/fr/products/firefox/>

-
- ¹⁴² <http://fr.techcrunch.com/2006/04/01/analysons-les-lecteurs-rss-en-ligne/>
<http://vtech.canalblog.com/archives/2005/10/25/927489.html>
<http://vtech.canalblog.com/archives/2005/09/05/780635.html>
<http://allrss.com/rssreaders.html>
- ¹⁴³ <http://influx.joueb.com/news/les-pages-personnalisables-bookmark-n-1>
http://www.informationoverlord.co.uk/2006/03/where-to-start-in-search-of-ideal_26
- ¹⁴⁴ <http://www.digimind.fr/produits/index.htm>
- ¹⁴⁵ <http://share.opml.org/>
- ¹⁴⁶ <http://www.quechoisir.org/Article.jsp?id=Ressources:Articles:EF33CF34D5616AF2C125720D0050E701&catcss=TEL000>
- ¹⁴⁷ <http://weblogs.hitwise.com/bill-tancer/>
- ¹⁴⁸ <http://technology.guardian.co.uk/weekly/story/0,,1823959,00.html?gusrc=rss>
- ¹⁴⁹ http://www.businessweek.com/innovate/NussbaumOnDesign/archives/2007/04/are_most_web_20.html
- ¹⁵⁰ http://blogs.forrester.com/charleneli/2007/04/forresters_new_.html
http://ross.typepad.com/blog/2006/04/power_law_of_pa.html
- ¹⁵¹ http://www.observatoire-omic.org/colloque-icic/pdf/Barbe3_1.pdf
- ¹⁵² <http://www.aaronsw.com/weblog/whowriteswikipedia>
- ¹⁵³ <http://aphgcaen.free.fr/chronique/393/aphg393.htm>
- ¹⁵⁴ http://fr.wikipedia.org/wiki/Droits_de_l%27homme_en_Iran
- ¹⁵⁵ http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Articles_de_qualit%C3%A9
- ¹⁵⁶ http://en.wikipedia.org/wiki/John_Seigenthaler_Sr._Wikipedia_biography_controversy
- ¹⁵⁷ <http://pewresearch.org/pubs/71/riding-the-waves-of-web-20>
- ¹⁵⁸ <http://tallblog.conted.ox.ac.uk/wp-content/uploads/2007/03/survey-summary.pdf>
- ¹⁵⁹ <http://www.cafepress.com/nerdwideweb>
- ¹⁶⁰ <http://www.ecrans.fr/spip.php?rubrique30>
- ¹⁶¹ <http://fr.creativecommons.org/>
- ¹⁶² <http://www.zdnet.fr/actualites/internet/0,39020774,39334479,00.htm>
- ¹⁶³ <http://websemantique.org/PagePrincipale>
<http://www.w3.org/DesignIssues/Semantic.html>
<http://www.w3.org/2002/Talks/1213-lp/slide1-0.html>
- ¹⁶⁴ <http://www.internetactu.net/?p=6902>